UNIVERSITY *of* HAWAI'I®

MĀNOA

Thuricin CD Biosynthetic Pathway: Isolation and Characterization of the Precursor Peptides and Sactisynthases TrnC and TrnD

Giacomo Mazzotti

Department of Chemistry

Table of contents

Figures and Tables

Abbreviations

5’-deoxyadenosine	5'-dAH
Acetonitrile	ACN
5’-deoxyadenosyl radical	5'-dA•
β-mercaptoethanol	β-ME
Chitin Binding Domain	CBD
Dithiothreitol	DTT
Electrospray ionization	ESI
Equivalents	eq./equiv.
Flavodoxin	Fld
Flavodoxin/ferredoxin:NADPH oxido-reductase from *T. elongatus*	FNR
High pressure liquid chromatography	HPLC
Isopropyl-β-D-1-thiogalactopyranoside	IPTG
Lysogeny broth	LB
Maltose Binding Protein	MBP
Mass spectrometry	MS
Molecular weight cut-off	MWCO
Nicotinamide adenine dinucleotide phosphate	NADPH
Nitrilotriacetic acid	NTA
Non-ribosomal polypeptide synthetase	NRPS
Optical density	OD
Polyacrylamide gel electrophoresis	PAGE

Ribosomally synthesized and post-translationally modified peptide	RiPP
Room temperature	r.t.
S-adenosyl-L-methionine	SAM
Sodium dodecyl sulfate	SDS
Sodium dithionite	DTN
Sporulation Killing Factor	SKF
Subtilosin A	SboA
Tobacco Etch Virus	TEV

Introduction

Ribosomally synthesized and post-translationally modified peptides

Antimicrobial peptides produced by all multicellular organisms as a first line of defense have often been shown an impact on human health[1]. Countless antibacterial, anticancer and antifungal therapeutics have been developed on the basis of these biologically active molecules[2]. Depending on whether their biosynthesis is ribosome dependent or independent, these natural products can be classified into two main categories[3]. Those produced by non-ribosomal polypeptide synthetases (NRPS) are the most broadly studied and include well-known antibiotics, such as vancomycin or bacitracin, which have been widely used over the past 50 years to treat different types of bacterial infections. The family of ribosomally-synthesized and post-translationally modified peptides (RiPPs) is in rapid expansion. Recent advancements in genome sequencing and bioinformatics have decreased costs, resulting in a vast increase to the amount of genomic data available. Due to the degree of complexity and diversity of most of these peptides, RiPPs were initially believed to be non-ribosomal peptides, which are assembled by large multi-modular enzymes. Their biosynthetic pathways are instead comparatively simple, and the structural diversity is introduced by extensive post-translational modifications. These modifications can be predicted with relative ease from the sequence homology of genes clustered around the peptide. The RiPP is initially synthesized as a short, unmodified peptide, together with several modifying enzymes and transporters to excrete the product after the modifications occurred[4]. The structure of the peptide operons is very similar across different organisms and share similar features that can be summarized. The key common component (Fig. 1) is the presence of a so-called "leader peptide" recognition sequence at the N-terminus or, more rarely, at the C-terminus ("follower peptide"). Typically, the biosynthesis starts with the synthesis of a longer

precursor peptide encoded by a structural gene, which includes the recognition sequence attached to the "core peptide", the linear and unmodified form of the RiPP. The core peptide is then post-

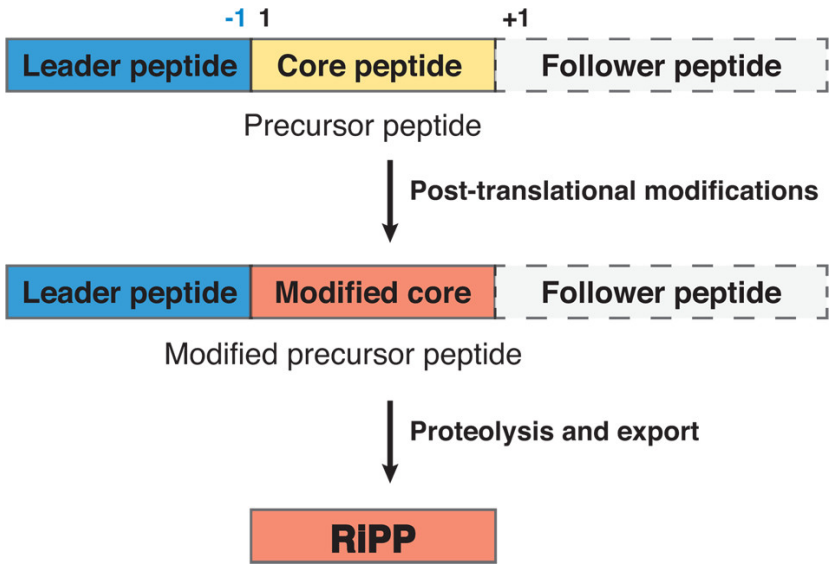


Figure 1. Ribosomally synthesized and post-translationally modified peptides general biosynthetic pathway.

translationally modified by a variable number of enzymatic reactions and usually after all the modifications are completed, the recognition sequence is proteolytically removed and the mature, bioactive form of the peptide is exported.

Overview of different structural features

Ribosomally-synthesized and post-translationally modified peptides with highly variable structures have been isolated from a variety of organisms. In reviewing the most interesting ones, it is worth mentioning the extremely cytotoxic polytheonamides, which belong to the proteusins family. Isolated from the marine sponge *Theonella swinhoei*, Polytheonamide A and B are complex peptides formed by 48 modified amino acids, with the peculiar presence of several *tert*-leucines, C-methylated amino acids and the almost perfect alternation of *L*- and *D*-configured

residues[5]. One of the remarkable aspects of their biosynthetic pathway is the presence of only six candidate enzymes in the entire 14kb gene cluster responsible for the 48 modifications (Fig. 2). Other examples of interesting structures from different organisms are the cyanobactins from cyanobacteria[6], N-to-C macrocyclic peptides usually produced when the bacteria are in symbiosis

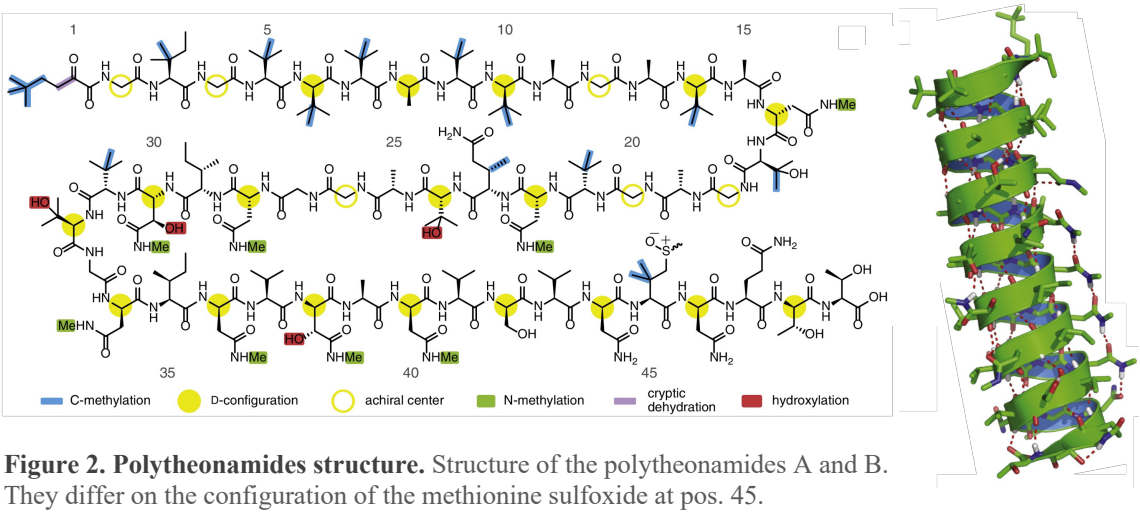


Figure 2. Polytheonamides structure. Structure of the polytheonamides A and B. They differ on the configuration of the methionine sulfoxide at pos. 45.

with another organism; the microcins produced by Enterobacteriaceae, very small peptides (<10 kDa), almost always encoded in plasmids[7]; the bottromycins found in common *Streptomyces*, containing a macrocyclic amidine and a decarboxylated C-terminal thiazole[8]. Very frequently, cysteine residues are involved in post-translational modifications, as thiols are readily converted to disulfides or thiazol(in)es for instance. Based on gene frequency in publicly available databases (NCBI, EBI, JGI, etc.), the most abundant RiPPs are in fact lanthipeptides, a distinct family of molecules with cysteine thioether crosslinks as their unique post-translational modification. To form this bond, dehydration of Ser or Thr leads to 2,3-didehydroalanine (Dha) or 2,3-didehydrobutyrine (Dhb) respectively, which can act as Micheal acceptors in a 1,4-conjugate addition with a cysteine residue nerby (Fig. 3A)[4]. Tautomerization of the resulting enolate completes the formation to a lanthionine (Lan, from Ser) or methyllanthionine (MeLan, from Thr) thioether crosslink. Depending on the enzymes responsible for the post-translational

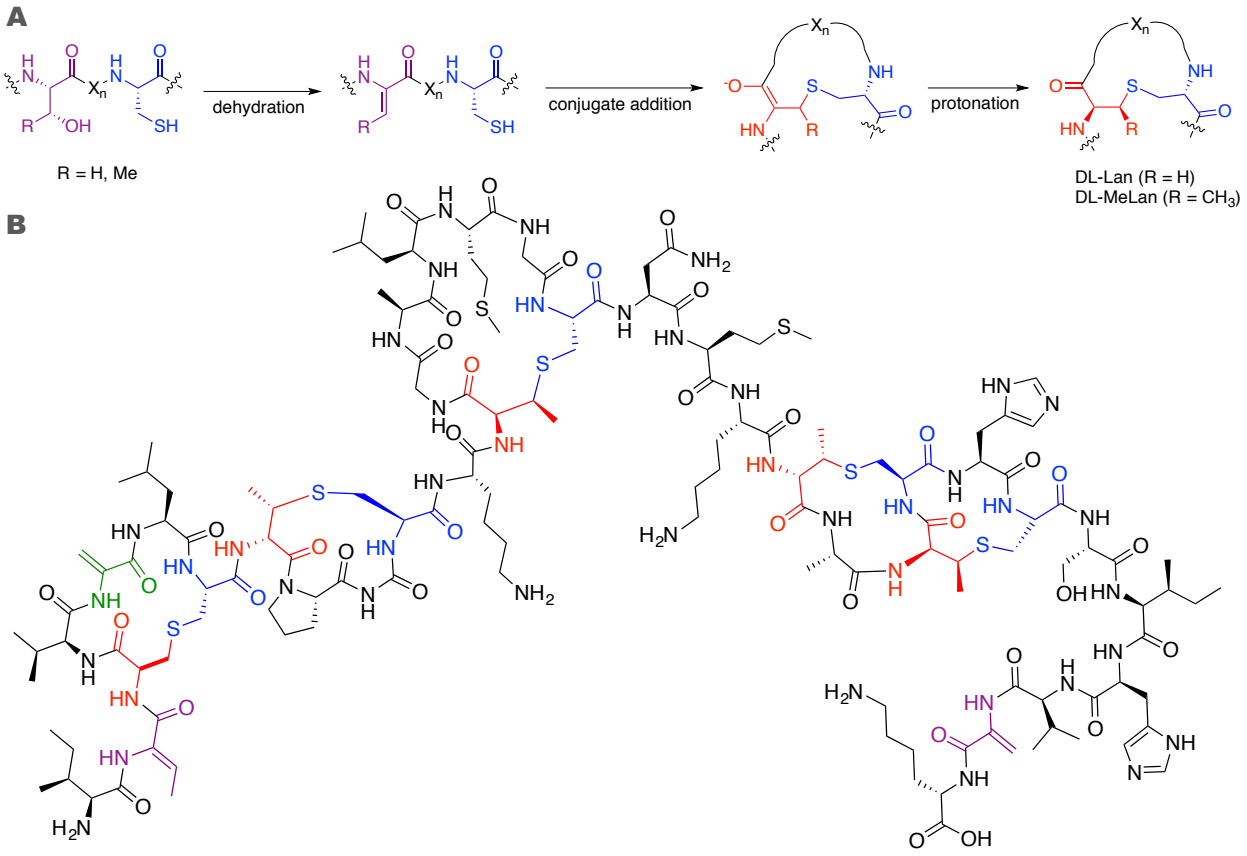


Figure 3. Lanthipeptides. A. Post-translational modification in lanthipeptides leading to the formation of (Me)Lan crosslinks. Phosphorylation and elimination lead to the formation of Dha or Dhb. **B.** Structure of nisin A. Dha residues are shown in green, Dhb purple, segments of (Me)Lan originating from Ser/Thr in red and segments from Cys in blue.

modifications, the lanthipeptides are divided in four classes (I-IV). Compounds that exhibit antimicrobial activity are classified as lanthibiotics. One of the first RiPP discovered was the lanthibiotic nisin, first isolated in 1928, and structurally fully elucidated in 1971[9]. Nisin A (Fig. 3B) contains five thioether crosslinks, each forming a ring of different size along the chain. Notably, thioether formation and ring closure result in re-protonation of the α-carbon at either the *re*- or *si*-face of the Lan or (Me)Lan residue, restoring or inverting the stereochemistry of the original amino acid. It is postulated that these stereochemical inversions, cyclizations or other modifications are common among all RiPPs, as they increase the metabolic stability of the peptide and decrease its conformational flexibility.

Sactipeptides

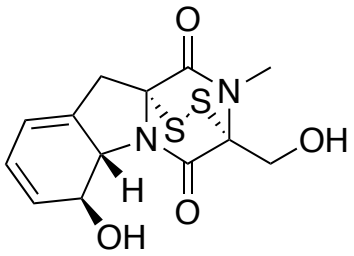


Figure 4. Structure of gliotoxin.

A different type of post-translational modification containing sulfur from cysteine residues is found in sactipeptides. This smaller group of more recently discovered RiPPs presents a characteristic crosslink between the sulfur of a Cys residue and the α-carbon of another residue. Given the strong antimicrobial activity and unifying structural features, the class has been named sactibiotics, from sulfur to alpha-carbon (or thioether cross-linked) antibiotics. This bond was already known to be present in some non-ribosomally synthesized diketopiperazine derivatives, such as gliotoxin (Fig. 4), formed by the nucleophilic attack of glutathione on a diketopiperazine *N*-acyl imine intermediate[10]. However, since it is NRPS-derived and involves a disulfide bridge, this modification differs from the new class introduced. To date, only seven sactipeptide natural products have been isolated or reported, six of which in *Bacillus* species and one in staphylococci. Subtilosin A (SboA) has been the first sactibiotic to be reported in 2003 by Vederas *et al*[11]. It is also the most well studied peptide of this family, as in the last 15 years several groups have shared efforts towards its understanding, from structure elucidation to the mechanism of action of its antimicrobial activity.

AlbA — Subtilosin A

Subtilosin A (Fig. 5) is encoded by the *sbo-alb* operon of *Bacillus subtilis*. It is initially ribosomally-synthesized as a linear 43-residue peptide and it undergoes several sequential post-translational modifications, before it is ultimately exported outside the cell membrane in its mature form, a circular head-to-tail 35-residue peptide. One of the key steps of its biosynthesis is carried

Figure 5. Subtilosin A. A. Structure of subtilosin A. In bold the leader peptide sequence, in red and green the residues involved in thioether crosslinks. **B.** Subtilosin A bioactive macrocycle exported outside the cell after several post-translational modifications.

out by the radical S-adenosylmethionine (SAM) enzyme AlbA, which catalyzes the formation of three sulfur to α-carbon thioether bonds on the initial linear moiety, between the cysteines C4, C7, C13 and three other residues F31, T28, F22 respectively[12]. The operon of each sactipeptide reported contains at least one *orf* encoding a radical SAM enzyme. These enzymes have been demonstrated to be responsible for the formation of the signature thioether bond, and they are therefore called sactisynthases. Radical SAM enzymes form a large superfamily containing over 100,000 homologous members that catalyze a number of different transformations, including metabolism, nucleic acid modification and biogenesis of cofactors[13].

The members of this family are characterized by the presence of a CXXXCXXC motif, where the three cysteine thiolates coordinate three iron atoms of a $[4Fe\text{-}4S]^{2+/+}$ cluster required for

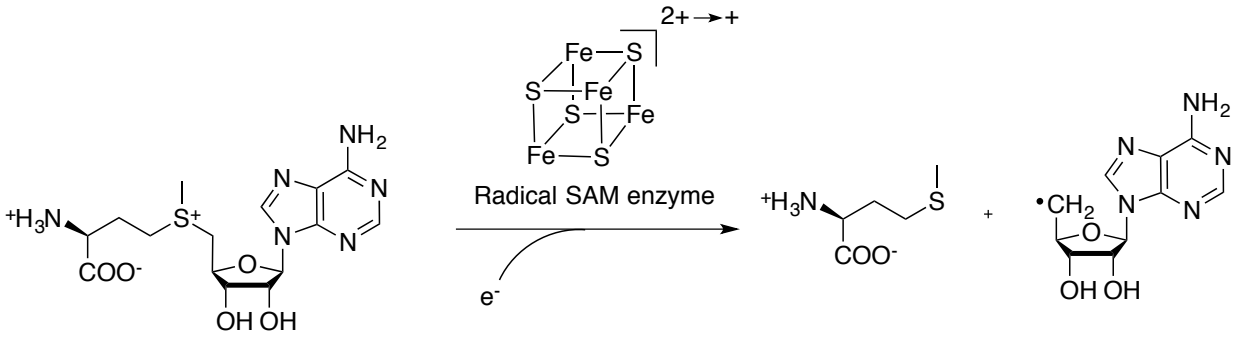


Figure 6. Radical SAM reaction scheme. General reaction scheme for radical S-adenosylmethionine enzymes. 5'-deoxyadenosine radical can now abstract a hydrogen from a non-reactive position.

catalytic activity. The remaining iron coordinates a molecule of S-adenosylmethionine, which, upon a single electron transfer from the reduced form of the Fe-S cluster ([4Fe-4S]$^{+}$), it cleaves generating 5'-deoxyadenosine radical (5'-dA•) and methionine, restoring the oxidized form of the cluster (Fig. 6). The cluster reduction from +2 to +1 requires low-potential electrons, which *in vitro* can often be successfully provided by NADPH/flavodoxin/ferredoxin systems or chemical reducing agents, such as sodium dithionite or titanium(III) citrate. The 5'-deoxyadenosine radical is then responsible for a variety of different reactions, typically activating aliphatic groups or other chemically inert positions. Moreover, for several radical SAM enzymes, the presence and function of auxiliary iron-sulfur clusters have been described. Two of the most remarkable examples of this

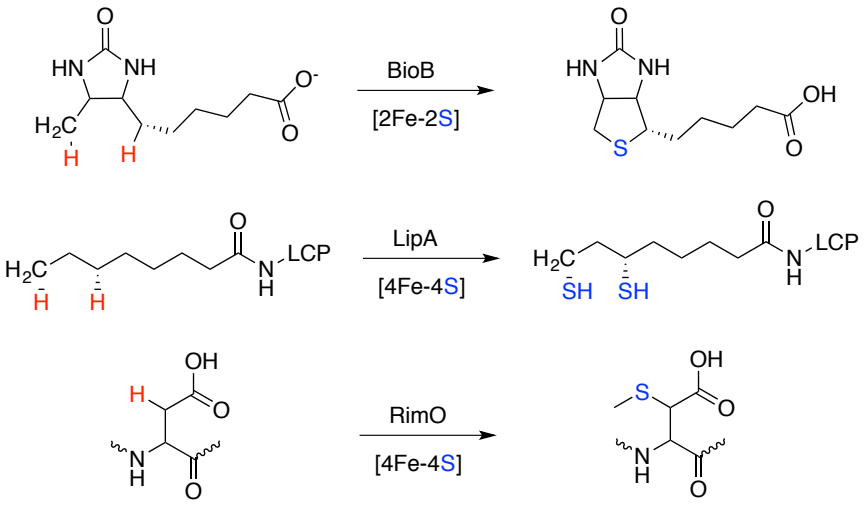


Figure 7. BioB, LipA, RimO. BioB, LipA and RimO reaction schemes. The auxiliary cluster they coordinate is indicated below the arrow. In red the hydrogens abstracted by 5'-dA•, in blue the sulfur atom from the Fe-S auxiliary cluster incorporated into the final product.

category are BioB[14], the enzyme responsible for the last reaction of the biotin biosynthetic pathway, which contains a [2Fe-2S]$^{2+/+}$ cluster in addition to the main [4Fe-4S]$^{2+/+}$, and LipA, involved in lipoic acid biosynthesis, which contains an auxiliary [4Fe-4S]$^{2+/+}$ cluster instead[15].

Following a similar mechanism, these two enzymes incorporate one (BioB) or two (LipA) sulfur atoms from the secondary cluster into their final product. For each sulfur to carbon bond, a hydrogen radical abstraction occurs, and thus one equivalent of SAM is required. LipA generates two sulfur-carbon bonds with two sulfur atoms, whereas BioB incorporates one sulfur atom into a thioether, closing a five-membered ring. Similar sulfur incorporation from auxiliary clusters is observed in methylthiotransferases, such as RimO and MiaB[16], where SAM first donates the methyl group to the sulfur of the cluster, which then reacts with the radical substrate to form a carbon-SCH_3 bond (Fig. 7).

In subtilosin A biosynthesis, AlbA contains two [4Fe-4S] clusters, with the second cluster most likely coordinated by a set of cysteines (C408, C414, C417) near the C-terminus. This domain is conserved within the subcategory of radical SAM enzymes classified as "SPASM and twitch domain-containing" maturases, where SPASM is an acronym for subtilosin, pyrroloquinoline quinone, anaerobic sulfatase and mycofactocin, the first natural products discovered to have a key step carried out by a protein from this family[17]. This domain is characterized by the presence of an additional variable cysteine motif near the C-terminus, usually involved in binding the auxiliary Fe-S

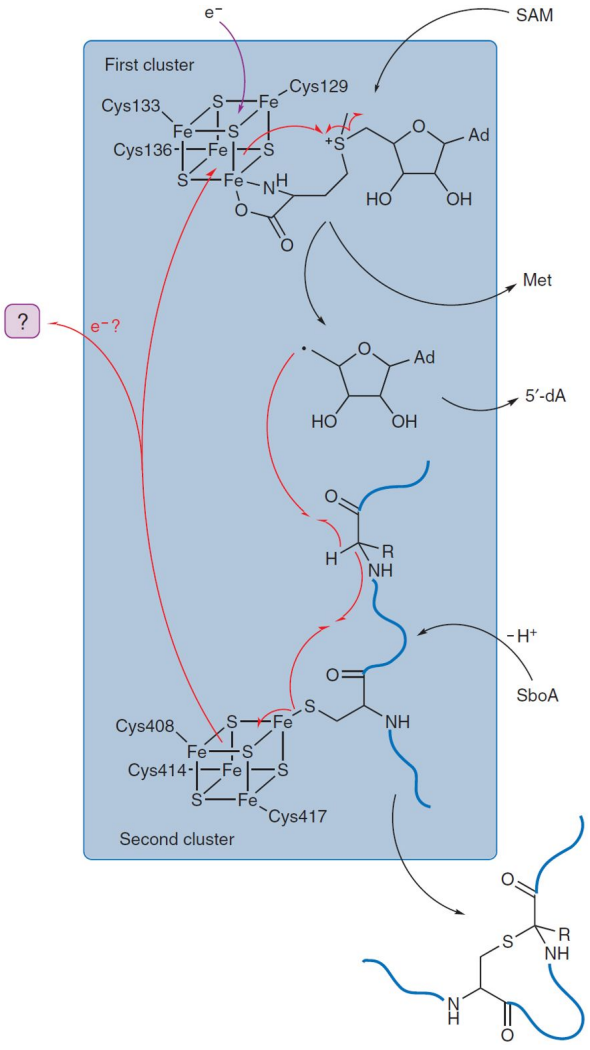


Figure 8. AlbA, proposed mechanism. Proposed mechanism for SboA thioether crosslink formation.

cluster(s). More specifically, a Twitch domain is a truncated version of the SPASM domain, containing no more than one auxiliary cluster.

For AlbA, in the sulfur to α-carbon crosslink formation mechanism, the substrate SboA is proposed to coordinate the auxiliary Fe-S cluster (Fig. 8). The first cluster would be responsible for SAM cleavage activity, forming a 5'-dA radical, which abstracts the α-hydrogen of a phenylalanine or threonine, and the carbon-centered radical would then form the thioether bond with the coordinated substrate cysteine, with the auxiliary Fe-S cluster acting as the electron acceptor.

ThnB — Thuricin H

The same research group proposed a similar mechanism for the radical SAM enzyme ThnB, in the biosynthesis of thuricin H[18,19], a 31-residue bacteriocin isolated from *Bacillus*

Figure 9. Thuricin H.

thuringiensis SF361. In this case, the precursor peptide ThnA undergoes formation of four different thioether crosslinks (Fig. 9). Interestingly, ThnB and AlbA share homologies in the N-terminal region with PqqD, an enzyme involved in pyrroloquinoline quinone cofactor (PQQ) biosynthesis. PqqD is proposed to be a peptide chaperone, forming a complex with the radical SAM enzyme PqqE, responsible for the formation of a C-C bond between the carbon in the *ortho* position to the phenol of a tyrosine residue and the γ-carbon of a glutamate residue. Marahiel's group has shown how a ThnB construct not containing this PqqD-like portion is still able to cleave S-adenosylmethionine, but not to catalyze the thioether crosslink. This led to proposing how, even without containing an active site, this region might be fundamental for substrate recognition and

binding, or for substrate positioning with respect to the second Fe-S cluster. Another hypothesis could be that the auxiliary clusters do not actively participate in the reaction, but they only move electrons from a reducing agent to the active site, which, in this case, would be the main radical SAM cluster. This mechanism is common for other proteins[20].

SkfB — Sporulation Killing Factor

The Sporulation Killing Factor (SKF) presents similar features to SboA and thuricin H. It is a 26-residue circular sactipeptide produced by *Bacillus subtilis 168*, which shows two post-translational modifications: a thioether linkage between Cys4 and Met12, and a disulfide bond between Cys1 and Cys16 (Fig. 10)[21]. The expression of its precursor SfkA is increased during limited nutrient conditions and, after modification and export from the cells, the peptide induces lysis or inhibits further growth of other microbial species. SfkB is the enzyme responsible for the formation of the C4-M12 thioether bond and it shows sequence homology with AlbA and ThnB, including the PqqD-like domain. The mechanism proposed for SfkB is equivalent to the enzymes containing auxiliary [4Fe-4S] described previously. This system is a good model for mechanistic studies of sactibiotic biosynthesis, as there is only one Cys to α-carbon crosslink formed. In addition to the experiments already mentioned, Marahiel *et al.* investigated SfkB specificity toward the donor and acceptor amino acid. The donor amino acid has to be a cysteine and does not react if exchanged into a serine, whereas the acceptor amino acid seems to be more flexible. More specifically, hydrophobic and aromatic amino acids are always tolerated as acceptors, small

Figure 10. Sporulation Killing Factor.

hydrophilic ones (Ser, Thr, Asp) can form the thioether bond in lower yields, and larger hydrophilic amino acids (Glu, Gln, Lys) are unreactive. This led to the conclusion that the radical α-hydrogen abstraction is independent of the amino acid nature, and only the second part of the mechanism dictates whether the cysteine will form the crosslink, or the radical substrate will abstract another hydrogen. Moreover, in all three enzymes AlbA, ThnB and SkfB, the thioether bond formation does not occur without the presence of the leader peptide sequence, suggesting that post-translational modification occurs prior to proteolysis. In 2018, Drennan *et al.* reported the crystal structure of SkfB[22], showing the secondary structure of the radical SAM domain, very similar to most enzymes of this family. In contrast to the hypothesis made by the Marahiel group, the crystal structure surprisingly shows the presence of a [2Fe-2S] cluster in the Twitch C-terminal domain (Fig. 11). Previously, PqqE had been shown to contain a [2Fe-2S] cluster[17], but SfkB is more homologous to members of the sactisynthase family, hypothesized to always contain [4Fe-4S] clusters. Similarly to PqqE, SfkB binds this auxiliary cluster into a $CX_{33}CXC$ motif, where the two

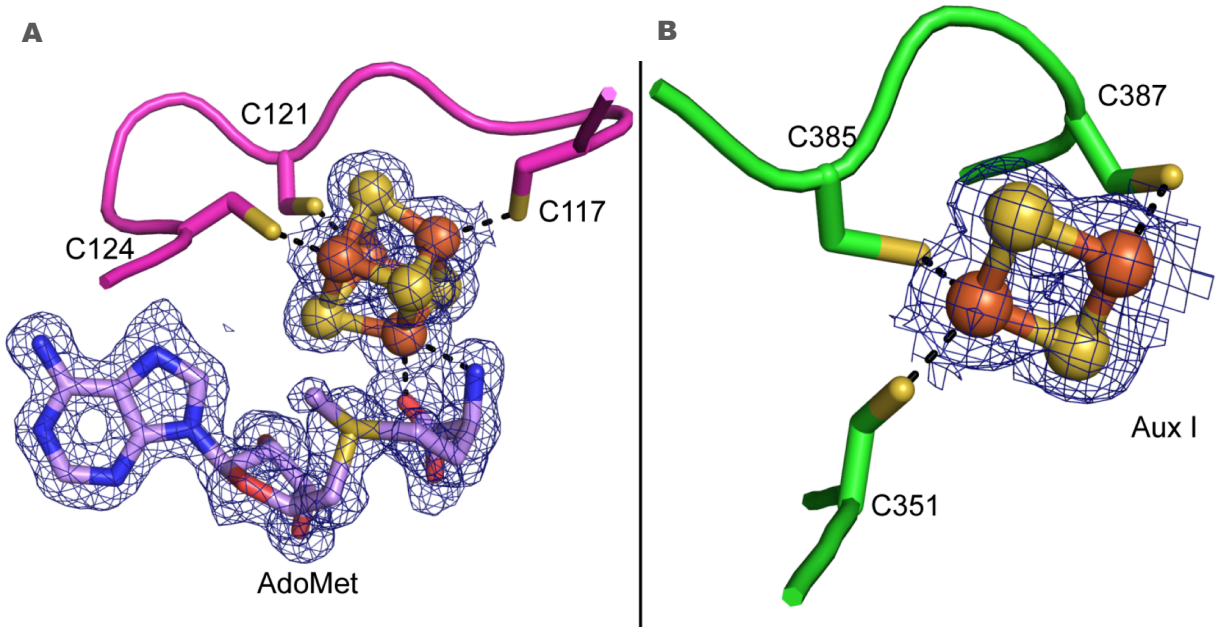


Figure 11. SkfB crystal structure. A. SkfB radical SAM domain [4Fe-4S] cluster; **B.** the [2Fe-2S] auxiliary cluster coordinated by SkfB Twitch domain.

nearly adjacent cysteine residues coordinate the same iron atom. Supporting the mechanism proposed for SboA, Drennan also discusses how this auxiliary [2Fe-2S] cluster could be potentially coordinated by the peptide and participate to the reaction, but, unfortunately, no crystal structure for a sactisynthase complex with its peptide has been reported.

Tte1186 — Tte1186a

An alternative mechanism has been proposed for the sactisynthase Tte1186 and its sactipeptide Tte1186a by Bandarian *et al*[23]. This enzyme is reported to contain three [4Fe-4S] clusters and, through mutation studies, it has been shown that in the absence of any of the two auxiliary clusters the reaction does not occur, even while cleavage of SAM does. The peptide contains six cysteines, but *in vitro*, the formation of only one crosslink was observed. Bandarian discusses the mechanism proposed by Marahiel, focusing on how the coordination of the secondary cluster to the substrate might be reasonable, but is not consistent with the stereochemistry of the final product. A majority of the sactipeptides that have been reported have multiple crosslinks, and the new carbon-sulfur bond forms with retention of stereochemistry at some centers and inversion at others. This stereochemical variability indicates that the radical intermediate likely can be attacked at either the *si*- or *re*-face. If this is the case, it has to be possible for the α-carbon radical to be attacked at the position initially occupied by the α-hydrogen (for crosslinks that occur with retention of configuration) or at the opposite face (in case of inversion). Assuming that the 5'-dAH does not dissociate from the active site until the end of the reaction, the face previously occupied by the α-hydrogen would be occluded by 5'-dAH, and a thiolate attack from this face would require the radical intermediate to be relatively flexible and rotate to expose this face. However, this option seems unreasonable, since the data available for radical SAM

enzymes shows how H-atom donor and acceptor are always within Van der Waals contact and movement of the radical intermediate is most likely limited and controlled by the active site.

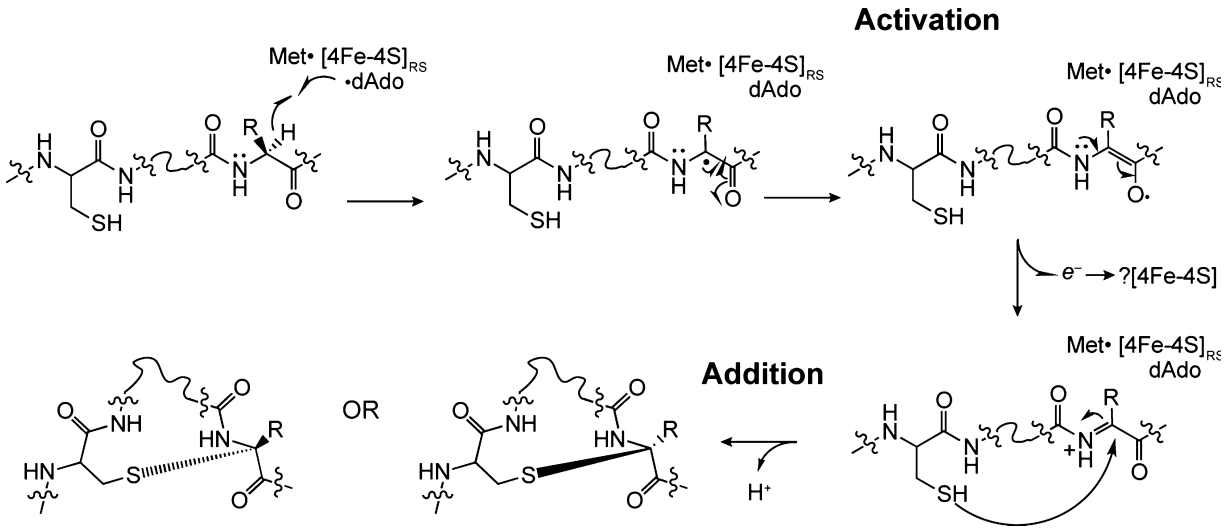


Figure 12. Thioether crosslink formation, mechanism proposed by Bandarian et al.

Therefore, Bandarian proposes not the direct formation of a thioether bond between Cys and α-carbon, but instead the conversion of the α-carbon radical to a more stable and planar intermediate, probably an acylimine or acylimminium, with the radical electron transferred to the auxiliary Fe-S cluster (Fig. 12). This intermediate could be attacked by the thiolate from either face. Bandarian's mechanism is analogous to the lanthionine bond formation in lanthipeptides[4], wherein an activated, planar amino acid intermediate is formed and regioselectivity is imposed by the enzyme through conformation of the active site.

Hyicin 4244 and thermocillin

Although they have not been studied as thoroughly as the other peptides described, two more examples have been recently added to the sactipeptides family and are worth mentioning. In 2018, Bastos *et al.* reported the bacteriocin Hyicin 4244 as the first sactibiotic produced by

staphylococci[24], making it the first produced by any organism other than a gram-positive spore-forming bacillus. The hyicin operon is very similar to those of subtilosin A and other members of the family, containing radical SAM enzyme, ABC transporter and peptidase genes among others. Unfortunately, the Bastos group has not been able to isolate the bacteriocin yet, nor any of the enzymes of the gene cluster. From gene sequence alignment, Hyicin 4244 is predicted to be very similar to subtilosin A structurally. Additionally, all the enzymes in its biosynthetic gene cluster show at least 42% sequence identity with the enzymes of the *sbo-alb* operon. This will be a very interesting molecule to isolate and study in the future, since concentrated biological samples showed the inhibition of biofilm formation in 14 staphylococci strains isolated from human infections and bovine mastitis. Hyicin 4244 has potential to become an antimicrobial for prevention or treatment of biofilm-related infections caused by several different staphylococci.

Moreover, thermocellin (cte) is a predicted natural product of the bacteriocins family from *Clostridium thermocellum ATCC 27405*, but it has not been successfully isolated yet. Bowers *et al.* have been able to heterologously express and purify its precursor peptide CteA and the sactisynthase CteB[25]. As for Tte1186a, CteA contains six cysteine residues, so presumably the natural product would contain six thioether crosslinks, since all sactipeptides reported so far are modified at each cysteine. However, the Bowers group was able to form only one thioether crosslink *in vitro*. The crystal structure of CteB shows the presence of two additional [4Fe-4S] clusters in the C-terminal domain, for a total of three clusters. Again, due to failed attempts in co-crystallization, there is no structure with full length CteA bound. Though, they were able to crystallize CteB with an N-terminal 21-residue truncation (CteA 1-21). According to the Bowers group, the binding position of this shorter peptide within CteB and the open coordination site of

one of the two auxiliary clusters imply that Marahiel *et al.* were right in proposing how the substrate coordinates to one of the auxiliary clusters.

Quinohemoprotein amine dehydrogenase γ-subunit

The final member of the sulfur to α-carbon crosslinks class is the small (~9kDa) quinohemoprotein amine dehydrogenase (QHNDH) γ-subunit QhpC[26]. This is a unique addition to the family, in that this protein is the subunit of a larger protein complex and not a peptide natural product. This implies that these thioether crosslinks are not just present in antimicrobial peptides, but they can be involved in other metabolic and constructive pathways as well. The radical SAM enzyme QhpD is responsible for the formation of three linkages in QhpC, on two Asp and one Glu residue. These bonds slightly differ from those of the sactipeptides previously described, in that they are sulfur to β-carbon crosslinks (Fig. 13), but the hydrogen homolytic abstraction step requires to overcome a higher bond dissociation energy: ~400 kJ/mol in QhpC vs. <350 kJ/mol for the sactipeptides described. QhpD coordinates three [4Fe-4S] clusters, all required for its activity. The mechanism proposed by Okajima *et al.* is equivalent to the one described for subtilosin A; however, in this case, QhpD forms a stable complex with QhpC (QhpD•QhpC). Mutation of any of the QhpD cysteine residue to Ser not only abolishes activity, but also drastically decreases the stability of the QhpD•QhpC complex. This raises the hypothesis of a role of the auxiliary [4Fe-4S] clusters in substrate binding.

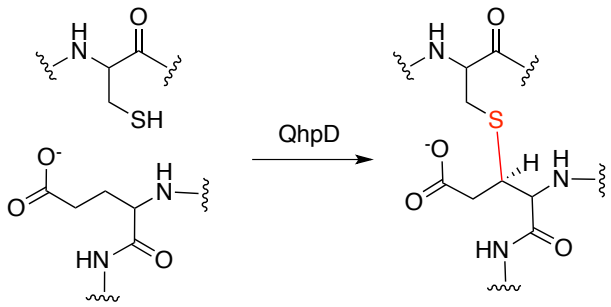


Figure 13. QhpD reaction scheme. Cysteine to β-carbon crosslink formation in QHNDH γ-subunit, mediated by QhpD.

Thuricin CD

One of the most well characterized sactibiotics to date is thuricin CD. Thuricin CD is a two-component bacteriocin, formed by two 30-residue subunits Trnα and Trnβ. These two peptides both contain three thioether crosslinks, specifically between Cys 5, 9, 13 and Thr 28, 25, Ser 21 respectively in Trnα, and Cys 5, 9, 13 and Tyr 28, Ala 25, Thr 21 in Trnβ, forming the sactibiotic secondary structure, which is most likely crucial for its bioactivity (Fig. 14). Thuricin CD was first isolated from *Bacillus thuringiensis* by Rea *et al.* in 2010[27], who, during the search for a potent narrow-spectrum antibiotic against *Clostridium difficile*, hypothesized that, in the presence of this bacterium, spore-forming anaerobic bacteria in the human gut would produce bacteriocins as a defense mechanism. Screening of over 30,000 fecal samples from healthy and diseased adults yielded only one colony with considerable activity. This colony was purified, grown in rich media and designated as *Bacillus thuringiensis DPC 6431*. Furthermore, the active natural product

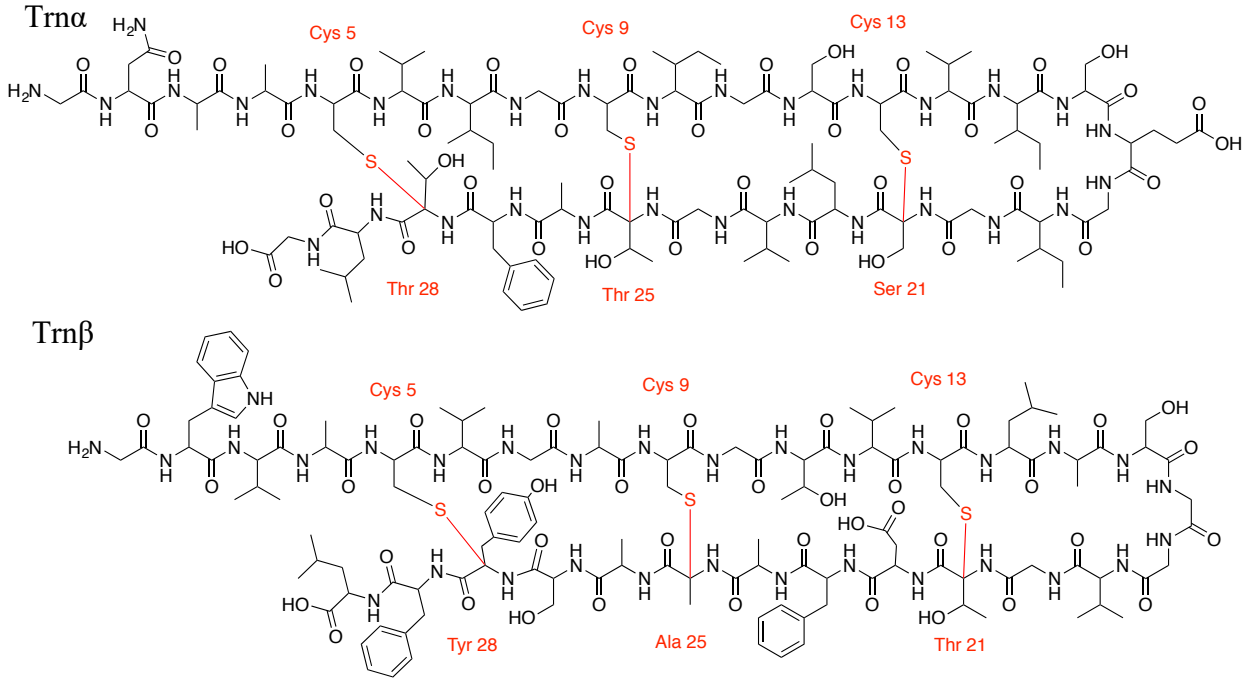


Figure 14. Thuricin CD. In red, the six cysteine to α-carbon crosslinks.

bacteriocin thuricin CD was isolated and characterized for the first time. The peptide shows little activity against almost all Gram-positive bacteria tested, but high potency against *C. difficile*, with an activity at least equivalent to commonly used antibiotics. This discovery makes thuricin CD a very strong candidate as a new narrow-spectrum drug against *C. difficile*.

Clostridium difficile and CDAD

C. difficile was defined as the causative agent of nosocomial diarrhea in 1978. This pathogen has become more common and dangerous each year globally, particularly with the appearance of a hypervirulent strain NAP1/BI/027, which developed in North America in the past decade[28]. It is strongly resistant to most common prescribed antibiotics, therefore it flourishes in the human gut as much of the beneficial flora is eradicated during wide-spectrum antibiotic therapy. It follows that, contracting an infection is easy in hospital environments, and immune compromised patients that come in contact with this pathogen will often develop *Clostridium difficile*-associated disease (CDAD). CDAD is an infection affecting almost half a million patients a year in the United States, with clinical illness ranging from mild diarrhea to fulminant colitis and death[29]. Around 30,000 deaths are associated to its diagnosis each year, with 80% of fatalities in patients over 65 years of age[30]. The transmission with this bacterium is primarily person to person, or by fecal-oral route and environmental contamination[31]; recent reports in Europe have also registered an increase in CDAD infections originating from animal reservoirs, such as the ribotype 087 strain of *C. difficile* (currently the second most common strain to cause human infections), which is often harbored in pigs[28]. The most commonly used antibiotics for CDAD treatment are metronidazole and vancomycin, both of which dramatically increase the abundance of organisms of the phylum Proteobacteria at the expense of organisms of other phyla[32]. Thuricin CD shows a

slightly weaker bioactivity compared to metronidazole, but it markedly improves the effect on the biodiversity of the gut microbiota (Fig. 15), making this peptide the most suitable antibiotic against

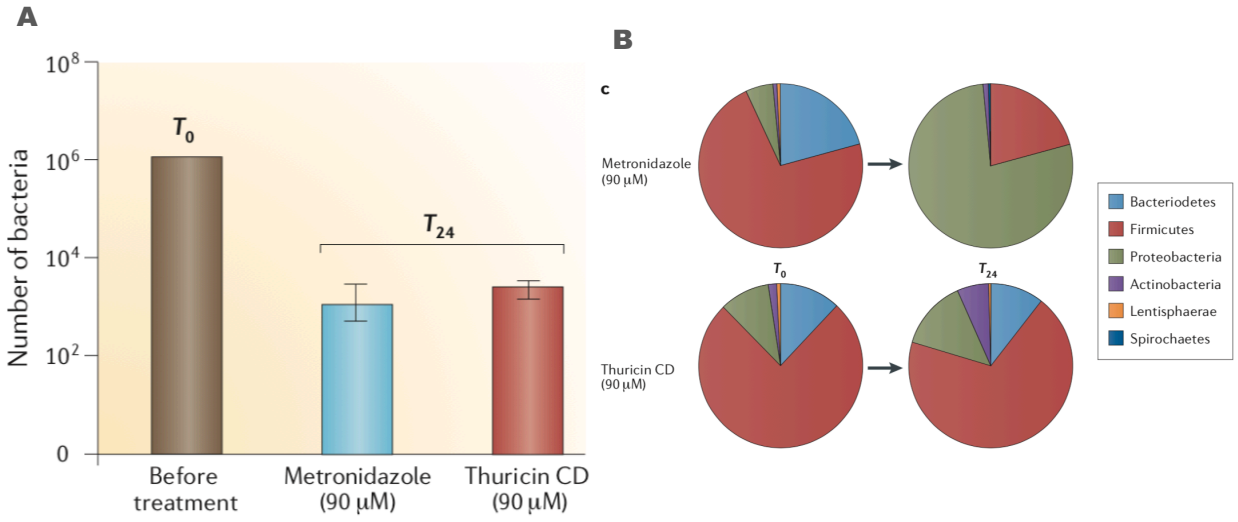


Figure 15. Metronidazole vs. thuricin CD. A. Comparison between metronidazole and thuricin CD after 24 hours incubation in a human colon infection model. **B.** Effect of metronidazole and thuricin CD on the composition of the gut microbiota, here depicted at the phylum level.

C. difficile known to date. A better understanding of its structure and biosynthetic pathway would inform efforts towards a biological method of production and aide in advancing clinical trials. The putative thuricin CD operon shows the same basic structure as subtilosin A and the other sactipeptides described (Fig. 16): two complete and overlapping *orf*s named *trnC* (1521 bp) and *trnD* (1245 bp) show homologies to the radical SAM superfamily and are thought to be responsible for the thioether crosslink formation; *trnF* and *trnG* genes are homologous to ATP binding cassette (ABC) transport proteins and together they are thought to form a translocation complex, with TrnF

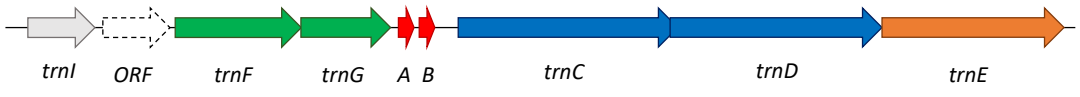


Figure 16. Thuricin CD gene cluster.

functioning as C-terminal ATP-binding domain, and TrnG as N-terminal hydrophobic integral membrane domain (since it includes six transmembrane α-helices)[33]. In 2014, Ross *et al.* identified an additional *orf* named *trnI* and located upstream of *trnF*[34], which encodes a small transmembrane protein that affords the producing microbe with resistance against thuricin CD; *trnE* shows homology to C-terminal processing peptidases; and finally, between *trnF/G* and *trnC/D* are located the two *orf*s *trnA* and *trnB*, encoding the precursor peptides to Trnα and Trnβ. TrnA and TrnB amino acid sequences are very similar, with a leader peptide of 17 and 19 residues respectively, followed by a Gly-Gly sequence and the core peptide, for a total of 47 and 49 residues (Fig. 17). Interestingly, the GG motif, which is commonly present in ribosomally synthesized bacteriocins, is usually cleaved by a protease after the second glycine. However, in thuricin CD the cleavage happens between the two residues. Assuming an equivalent biosynthetic pathway to the other members of the sactipeptide family, thuricin CD is proposed to be synthesized as a linear precursor peptide, which undergoes thioether crosslink formation by TrnC and TrnD as the first post-translational modification, followed by cleavage of the leader sequence by TrnE peptidase. The subsequent transport and excretion of the mature peptide is mediated by TrnF and TrnG (Fig. 18). Vederas *et al.* reported multidimensional NMR studies, isolating thuricin CD from fermented *B. thuringiensis* in [^{13}C, ^{15}N] fully labeled rich media[35]. Through NMR studies and calculation of the most optimal stereoisomers, the authors determined the stereochemistry of the amino acids involved in the thioether crosslinks. Interestingly, both peptides show an R-configuration in two of the crosslinks and a S-configuration in the third one, Thr28 in Trnα and Tyr28 in Trnβ. Since thuricin CD is the first case of a sactipeptide with two *orf*s for two separate radical SAM enzymes in its gene cluster, this finding raises several questions about the mechanism of formation of thioether bond formation, the order in which they form, the specific role of each enzyme and how

thuricin CD biosynthesis differs from the other members of its family. When target bacteria are treated with a single peptide subunit, Trnβ is 10 times more active against *C. difficile* than Trnα. However, when Trnβ and Trnα are combined at an optimized ratio of 2:1 (forming thuricin CD), the bioactivity increases 100-fold, reaching a minimum inhibitory concentration $(MIC)_{50}$ of 0.05 μM, showing how the two peptides act in synergy.

TrnA (47 residues): MEVMNNALITKVDEEIGGNAACVIGCIGSCVISEGIGSLVGTAFTLG

TrnB (49 residues): MEVLNKQNVNIIPESEEVGGWVACVGACGTVCLASGGVGTEFAAASYFL

Trnα: GNAACVIGCIGSCVISEGIGSLVGTAFTLG

Trnβ: GWVACVGACGTVCLASGGVGTEFAAASYFL

Figure 17. Amino acid sequences of TrnA, TrnB, Trnα and Trnβ.

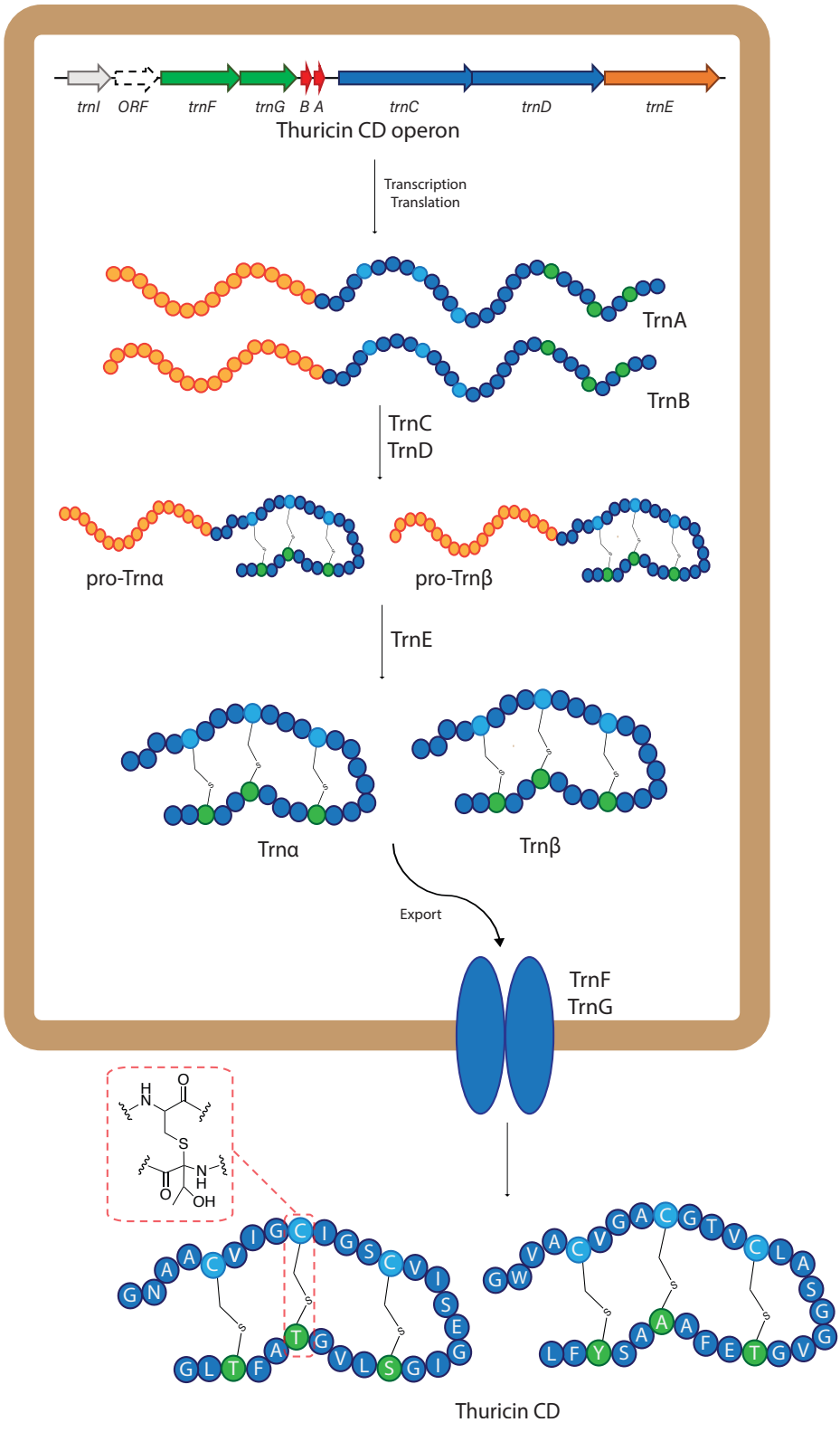


Figure 18. Thuricin CD, proposed biosynthetic pathway.

TrnC and TrnD

Despite having general features similar to the other sactisynthases mentioned, with a CXXCXXXC motif at the N-terminus and the presence of several other cysteines near the C-terminus SPASM domain, both TrnC and TrnD only show less than 20% sequence identity to AlbA. TrnC also shows sequence similarity to conserved domains of Fe-S oxidoreductase arylsulfatase regulatory proteins (AslB), whereas TrnD sequence is similar to enzymes in the MoaA family, which are involved in molybdopterin biosynthesis[27]. Given the stereochemical inversion on the last amino acid crosslink in Trnα and Trnβ, it is possible that one of these two enzymes is an epimerase and perhaps the other responsible for all three crosslinks. Alternatively, one could imagine that one enzyme generates two crosslinks and the third crosslink is installed by the second sactisynthase, possibly through a different mechanism, with inversion of the stereochemistry. The epimerase option is valuable, since post-translational epimerization generated by radical SAM epimerases in *Bacillus* species has been reported before[36].

Alternatively, since *trnC* and *trnD* genes are overlapping, another hypothesis is that the two proteins form an active complex (TrnC-D), wherein the separate proteins would not bind the substrate (TrnA or TrnB) until they assume a certain conformation via dimerization. This phenomenon is observed in the proteins PqqD and PqqE in the pyrroloquinoline quinone

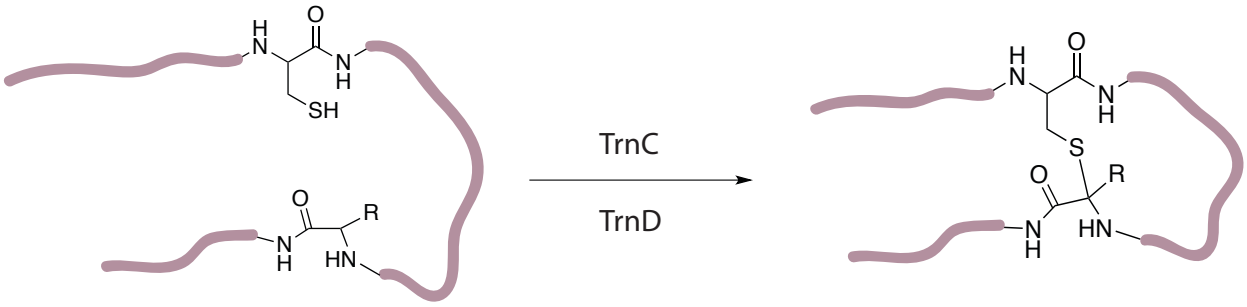


Figure 19. TrnC and TrnD general reaction scheme.

biosynthetis pathway. Here, PqqD plays a role as peptide chaperon binding the peptide PqqA. This event then triggers the radical SAM enzyme PqqE to bind the PqqD-A complex and catalyze a C-C crosslink formation[37].

TrnC (516 residues):

MSKKRLSMMPFKTDKFSYALDGNTGRVIVADKPTLYIISHFHKFEKEELLKKTKGKFAELHQDYLTTYNYVSSLINM
GMFYLSEKEDSDSPIDSKELAINSNQSQLILILTEKCNLRCEYCIYNDKYPKEMGYSDEEMDFETAKKAVDMYYELH
MERVKRGHKRFPVITMYGGEPLLKFDLIKKVMEYAKGLMPDTLFYTTTNGTLLSEKMMDYFINNRIIITFSIDGFKE
NHDRNRVFVNGMPTFERAFKNIKRLQEKKKEQNIEQIISFNCCFDQYTDVYKVAKFFEEHYDLFNPFFVLYNQINPF
DTLYFDWCDEQVKTGKWNFDKNNFKNAMQKIEHELYEAETCDDHFQQVAGPLVMKDFVLSIRNKDGQQQITRNSCIP
TSKIAVSPDGTLTLCEKMCKKYPIGTVEKGLDWKAVDGVTEKLVRHFNSDSCKYCPIRTMCEACFMFLDENGRIKPS
FCKSKKMAVKKNLESYFAKKEKGFDVMKVYNHTSDLDSVKEMVK

TrnD (414 residues):

MKKYFRLYPYCHLEIGETNSCLYDISSGKMIRVNRENAELLRQCQENVPIESINMDLGILDELIKMNLGTYYANPQF
IEPFFETNDTKNRIFGKNNILRQMFILTSTDCNMNCKHCNTDSTVFRKTGCKIWPKSINLNALTQSHWRKILEAFYN
LHGEELTFIGGEPFLEFDFIKNIVEIAQEVGISKFSIFTNGSIINDTILNFLMENKIKVYIQIFEVDENKFKAFTNS
DIPSIQIIDNIKKLNNHHLDLQLRILITRDNDNNLKKIVNTLQKETNVKDIKIEFLYPKPDNSYYSKKYIPLMYDKK
REFSHVNVQKMQFLHQYNPSFFGQITIRRDGKVVPHPMLLTRVIGDLQQDDLFTIINTEEYQEYSTLNKEKISKCST
CAYKYNCMDDRVIENFATGDLYGMEYCNF

Figure 20. TrnC and TrnD sequences. In blue, the CXXXCXXC motif. In red, other Cys present in the sequence. Near the C-terminus the Cys indicated are part of the SPASM domain and they might be coordinating auxiliary Fe-S cluster(s).

Research Statement

Considering the potential of thuricin CD as a drug, the aim of this project is to understand its biosynthesis, with the ultimate goal of producing the bacteriocin in large scale. To date, the mature bioactive peptide has been thoroughly characterized, but none of the enzymes responsible for the post-translational modifications have been studied at all. The two subunits forming the active peptide thuricin CD, Trnα and Trnβ, each feature three cysteine to α-carbon crosslinks. This unique structure is the hallmark of the sactipeptides family, a class of RiPP natural products counting only a few known members. These thioether crosslinks are known to be formed by radical S-adenosylmethionine enzymes, and while several mechanisms for installation of these bonds have been proposed, little enzymology data exists to support them. Usually, the sactipeptide operons contain one of these enzymes, which is responsible for one or multiple crosslinks. Thuricin CD is comprised of two different peptide subunits that works in synergy, and the gene cluster shows the presence of two radical SAM enzymes, whose specific functions are unknown. Heterologous expression in *Escherichia coli*, purification and initial characterization of TrnC and TrnD will be presented, in addition to methodology for the expression and purification of their substrates TrnA and TrnB, by way of a fusion protein system. Attempts to the understanding of the system will also be discussed, laying the groundwork for further investigations.

Materials and Methods

TrnC/TrnD cloning

Two pUC57 vector containing *trnC* and *trnD* genes were purchased and the two genes were cloned out and ligated into pET-28a(+) plasmids with traditional cloning procedures. The plasmids were transformed into *E. coli* BL21 Star (DE3) cells using the heat shock protocol.

TrnC/TrnD expression

The transformed cells were grown in lysogeny broth medium (LB, Fisher Scientific) containing 50 μg/mL kanamycin, at 37°C, 225 rpm for 12-16 hrs. The cells were then transferred into culture flasks containing 1 L LB, 50 μg/mL kanamycin and 5 mM ferric citrate and let grow at 37°C, 225 rpm. At an optical density (OD) of 0.5 (600 nm), the temperature was reduced to 22°C and isopropyl-β-D-thiogalactopyranoside (IPTG) was added to a final concentration of 0.2 mM. The cells were grown for 5-6 hrs, then harvested at 5,000 rpm for 10 min and frozen at -80°C.

TrnC purification

The cell pellets were resuspended in Buffer A (50 mM Tris/HCl, 0.5 M NaCl, 20 mM Imidazole, pH 8.0), transferred into a glass beaker and lysed using an ultrasonic homogenizer (two cycles, 2 min process time, 30 sec ON, 2 min OFF), then transferred into high speed centrifuge tubes. The lysate was centrifuged at 16,000 rpm for 1 h 10 min, the supernatant was kept, quickly transferred into a different tube and moved inside an inert atmosphere glove box, to minimize the oxygen exposure. The tube was opened inside the glove box and let incubate under nitrogen for 2 hrs. TrnC was purified through gravity-flow Nickel-chelate chromatography (Ni-NTA), using

Qiagen Ni-NTA agarose resin (Qiagen). The column was equilibrated with 8 column volumes of Buffer A, the lysate containing TrnC loaded and the column washed thoroughly with Buffer A (>10 column volumes). A dark brown band was observed on top of the resin. The protein was eluted with Buffer B (50 mM Tris/HCl, 0.5 M NaCl, 300 mM Imidazole, pH 8.0) and the darkest fractions pooled together. Purification proceeded with iron-sulfur cluster(s) reconstitution.

TrnD purification

The cell pellets were resuspended in Buffer A (50 mM Tris/HCl, 0.5 M NaCl, 20 mM Imidazole, pH 8.0), transferred into a glass beaker, lysed and centrifuged using the same procedure described for TrnC. Since TrnD is more air-stable than TrnC, the purification process could be carried out in normal atmosphere. TrnD was purified by FPLC, using Ni-NTA agarose resin. The column was equilibrated at 5 mL/min with Buffer A, the lysate containing TrnD was then loaded at 2 mL/min and the column washed with Buffer A at 2 mL/min until the absorbance was stable at ~0.2 AU. A dark brown band was observed on top of the resin. The protein was eluted with Buffer B (50 mM Tris/HCl, 0.5 M NaCl, 300 mM imidazole, pH 8.0) at 5 mL/min and the collected fractions concentrated using Amicon Ultra Centrifuge Filters 30k MWCO (Millipore). TrnD was then either reconstituted using the procedure described for TrnC or desalted through a P2 size-exclusion resin (Bio-RAD) (elution buffer: 50 mM Tris/HCl, 5 mM βME, 20% glycerol, pH 8.0).

Fe-S cluster(s) reconstitution

For longer storage time and iron and sulfur content analysis, TrnC and TrnD Fe-S clusters were reconstituted after Ni-NTA affinity chromatography. All required solutions (1 M dithiothreitol (DTT), 100 mM sodium sulfide (Na_2S) and 100 mM ammonium iron citrate) were

freshly prepared with reconstitution buffer (25 mM Tris/HCl, 150 mM NaCl, 10 mM DTT, pH 8.0). One hundred equivalents DTT were added to 50 μM TrnC(D) and incubated at r.t. for 15 minutes, followed by slow addition of 10 eq. of ammonium iron citrate, further 5 min incubation at r.t. and final addition of 10 eq. of Na_2S. The resulting dark brown solution was incubated on ice overnight, then purified by gravity-flow size-exclusion chromatography, using Sephadex G-25 Medium resin (GE Healthcare). Reconstitution buffer was used as final elution buffer, the darkest fractions were further concentrated down to the desired concentration, using Pierce PES concentrators, 10k MWCO (Thermo) and small aliquots frozen at -80°C, until further use. TrnC and TrnD were identified by sodium sulfate-polyacrylamide gel electrophoresis (SDS-PAGE).

Iron and sulfur content determination

The iron content in TrnC and TrnD was determined following a modified version of Collins *et al.* analytical method[38]. A calibration curve was prepared from a 1000 mg/L iron standard stock solution (Sigma), with standards ranging from 5 to 25 nmol Fe. Considering the possible presence of 2 or 3 [4Fe-4S] clusters, the samples were prepared aiming for ~15 – 20 nmol Fe. Samples and standards were diluted to a volume of 500 μL with Milli-Q H_2O in 1.5 mL Eppendorf tubes and followed by addition of 450 μL of 1 M HCl/10% trichloroacetic acid in H_2O. The solutions were vortexed, incubated at 37°C for 30 min and any precipitate removed by centrifugation. The supernatant was transferred to clean tubes and 100 μL of 10% (w/v) hydroxylamine•HCl were added; the samples were vortexed and 200 μL of 4 mM 2,4,6-tripyridyl-*s*-triazine in H_2O were added, followed by 150 μL of 50% ammonium acetate. The samples were centrifuged at 14,000 rpm for 10 min to remove any precipitate and the absorbance of each sample was read at 596 nm (ε_{596} = 22,600 M^{-1} cm^{-1}).

Sulfur was determined using Beinert's method for iron-sulfur proteins[39]. A calibration curve was prepared using $Na_2S•9H_2O$, ranging from 2 to 50 nmol S. Again, considering 2 or 3 [4Fe-4S] clusters, the samples were prepared aiming for ~20 – 30 nmol S. Samples and standards were diluted with H_2O to a final volume of 200 μL in a 1.5 mL Eppendorf tube, followed by addition of 600 μL of 1% (w/v) zinc acetate and 50 μL of 7% (w/v) sodium hydroxide. The solutions were mixed by inversion, incubated at room temperature for 15 min and the precipitate centrifuged at 3,000 rpm for 10 s. In the same tube, 150 μL N-N'-dimethylbenzene-1,4-diamine (DMPD) (0.1% (w/v) in 5 M HCl) and 150 μL $FeCl_3$ (10 mM in 1M HCl) were added and the pellets resuspended by vortexing. The samples were incubated at r.t. for 20 min to allow the formation of methylene blue and the absorbance was measured at 670 nm (ε_{596} = 71,089 M^{-1} cm^{-1}).

Iron-sulfur clusters qualitative analysis

In an inert glove box, the two proteins were transferred into an air tight cuvette at a concentration of 25-30 μM. A full absorbance spectrum from 200 to 800 nm was recorded (Varian Cary 50 spectrophotometer). After addition of 100 equiv. of sodium dithionite, another full scan was recorded, and the two spectra compared.

Titanium(III) citrate preparation

Titanium(III) citrate was prepared according to Ghabbour[40] procedure. A sure-sealed bottle of 10% $TiCl_3$ in HCl (Sigma) was carefully moved into an inert atmosphere glovebox. In a round-bottom flask, 500 μL of $TiCl_3$ were slowly added to 10 mL of a 0.2 M sodium citrate solution with continuous stirring. Sodium hydroxide was added dropwise until pH reached 8.0 and the solution

was used within a couple hours because prompt to fast oxidation. Titanium(III) citrate concentration was estimated to be ~80 mM.

S-Adenosylmethionine cleavage assays

The S-Adenosylmethionine (SAM) cleavage assays for TrnC and TrnD were carried out under anaerobic conditions in an inert atmosphere glovebox. All the stock solutions were brought into the box and incubated under nitrogen for about 1 hr. Three reducing systems were tested: 50 μM Flavodoxin (Fld)/5 μM Ferredoxin (FNR)/1 mM NADPH, 300 μM sodium dithionite (DTN) or 10 mM titanium(III) citrate. The final volume for each assay was 200 μL and the buffer used was 50 mM Tris/HCl, 0.1 M NaCl (pH 8.0). In order to further reconstitute the iron-sulfur clusters, 25 μM of TrnC or TrnD were incubated at r.t. for 15 min with 5 mM DTT, 250 μM ammonium iron(III) citrate and 250 μM sodium sulfide. In the case of Fld/FNR/NADPH reducing system, the two proteins were added with the reconstitution buffer and NADPH was added with SAM. The mixture was then incubated with 300 μM SAM at r.t. for 15 min and finally, DTN and titanium citrate were added. Controls with no SAM, or no enzyme, or no reducing system were also carried out. The assays were run for 3-4 hrs, the proteins denatured with 10% sodium acetate pH 4.5, the reaction tubes placed on ice for 15-20 min and the precipitate spun down in a microcentrifuge at 14,000 rpm for 10 min. The supernatant was analyzed by HPLC-MS. Chromatographic separation was performed on a T_5-C18 column (Waters, 4.6 mm X 150 mm, 5 μm) at 0.4 mL/min. The column was equilibrated in buffer A (1 mM NH_4HCO_2 in H_2O, pH 3.5), and upon injection of 10 μL of the sample, the reaction components were eluted with a linear gradient 0-20% buffer B (1 mM NH_4HCO_2 in 95:5 acetonitrile/H_2O, pH 3.5) over 15 min. The assays spectra were quantified at 260 nm and compared to a five point calibration curve of analytical standards (Sigma), 5'-

deoxyadenosine (t_R = 11.0 min) ($[M+H]^+$ = 252.1). The ions were analyzed by an Agilent 6400 triple-quadrupole mass spectrometer, equipped with an electrospray ionization (ESI) source, in positive mode. Nitrogen was the nebulizer drying gas (press. = 30 psi; flow rate = 7 L/min; temp. = 350°C). Additional parameters: spray capillary voltage = 3500 V; skimmer voltage = 65 V; fragmenter voltage = 100 V. Spectra were acquired over a 100-400 m/z range at 0.99 scan/s.

MBP-TrnB-CBD cloning

The vector pTWIN1 encoding the chitin binding domain (CBD) was purchased (NEB). A set of primers was designed in order to clone out of the plasmid the CBD gene, and another set to insert it at TrnB C-terminus, into the pMAL-c2x TrnB plasmid. After successful cloning, the pMAL-c2x TrnB-CBD plasmid was transformed into *E. coli* Rosetta (DE3) competent cells.

MBP-TrnB and MBP-KR5/KR10-TrnB cloning

A pUC57 vector containing *trnA* and *trnB* genes was purchased. The gene *trnB* was amplified and ligated following the manufacturer's protocol into a modified pMAL-c2x plasmid (NEB), with a TEV protease cleavage site instead of the commercial Factor-Xa one. The plasmid was transformed into *E. coli* BL21 Star (DE3) cells using the heat shock method. The MBP-TrnB plasmid was also used as a template for KR5. The cloning was carried out aiming to introduce an additional five lysine/arginine residues at TrnB N-terminus. The MBP-KR5-TrnB plasmid was used as a template for KR10, with the purpose of introducing five more Lys/Arg. Every plasmid described was confirmed by agarose gel analysis and DNA sequencing transformed by heat shock into BL21 Star (DE3) cells.

MBP-TrnB, MBP-TrnB-CBD and MBP-TrnB/KR5/KR10 expression

The transformed cells were grown in LB medium containing 50 μg/mL ampicillin, at 37°C, 225 rpm for 12-16 hrs. The cells were then transferred into culture flasks containing 1 L LB and 50 μg/mL ampicillin and let grow at 37°C, 225 rpm. At an optical density (OD) of 0.5 (600 nm), 0.2 mM IPTG was added and the cells let express for 4 hrs at 37°C, 225 rpm, then harvested at 5,000 rpm for 10 min and frozen at -80°C.

MBP-TrnB, MBP-TrnB/KR5/KR10 and MBP-TrnB-CBD purification

The cell pellets were resuspended in Buffer C (20 mM Tris/HCl, 0.2 M NaCl, 1 mM EDTA, 1 mM DTT, pH 7.4), transferred into a glass beaker, lysed and centrifuged using the same procedure described for TrnC. MBP-TrnB was purified by FPLC, with an amylose resin affinity column (NEB). The column was equilibrated at 5 mL/min with Buffer C, the lysate containing MBP-TrnB was then loaded at 2 mL/min and the column washed with Buffer C at 2 mL/min until the absorbance was stable at ~0.2 AU. The protein was eluted with Buffer D (Buffer C + 10 mM maltose) at 5 mL/min and the collected fractions concentrated using Amicon Centrifuge Filters 30k MWCO (Millipore). MBP-TrnB desalted through a P2 size exclusion resin (Bio-RAD) (elution buffer: 50 mM Tris/HCl, pH 8.0), concentrated again, identified by SDS-PAGE using PageRuler Unstained Protein Ladder (Thermo) and stored at -80°C.

Cleavage of CBD fusion

Cleavage of the chitin binding domain was carried out following a modified protocol from the one provided by the manufacturer (NEB)[41]. After the initial amylose purification and desalting steps, the protein solution was adjusted to 20 mM Tris/HCl (pH 8.5), 0.5 M NaCl, 1mM EDTA.

The cleavage reaction was initiated in solution adding 40 mM DTT and the mixture was incubated at 4°C overnight. A chitin resin column was equilibrated with the same buffer and the cleavage mixture was loaded. The flow-through was collected and CBD was eluted with 0.3 M sodium hydroxide. The two fractions were evaluated by SDS-PAGE. The resulting protein MBP-TrnB was dialyzed overnight in 50 mM Tris/HCl (pH 8.0).

Cleavage of MBP fusion

Tobacco Etch Virus protease gene was purchased in a pRK793 vector, transformed in BL21 Star (DE3), expressed and purified using the procedure described for TrnD. TEV protease was divided in small aliquots and stored at -80°C until further use. Cleavage of the MBP portion by TEV protease was carried out following a standard procedure[42]. The protein solution was adjusted to 50 mM Tris/HCl (pH 8.0), 0.5 mM EDTA, 1 mM DTT. TEV protease was added in a ratio of 1:100 protease/target protein and the cleavage carried out at 4°C overnight. A Ni-NTA resin column was equilibrated with the cleavage buffer. The crude mixture was then loaded onto the column and the flow-through collected in fractions. The cleavage buffer with the addition of 200 mM imidazole was used to elute TEV and MBP bound to the column. The flow-through fractions and the final elution were concentrated using 3k and 10k MWCO Amicon spin filters respectively, and analyzed by SDS-PAGE and HPLC-MS. The analysis was carried out in the same LC-QQQ system described for the SAM cleavage assays. Chromatography was performed on an Eclipse XDB C18 column (Agilent), flow 0.7 mL/min, with a linear gradient 20-80% acetonitrile in H_2O (with 0.1% formic acid. The UV trace was recorded at 210 nm.

SUMO-TrnA/TrnB cloning

Two micrograms of pETHSUL vector were treated with 4 units of BseRI and 4 units of FastAP thermosensitive alkaline phosphatase (Thermo) in a 50 μL reaction tube, with CutSmart buffer (NEB). The mixture was incubated at 37°C for 15 minutes, then placed on ice. Since the restriction digest was not completed, the linearized plasmid was purified by agarose gel electrophoresis and Illustra GFX PCR DNA and Gel Band Purification Kit (GE Healthcare). Using a thermal cycler, 0.1 pmol of the eluted DNA were incubated for 25 minutes at 22°C with 2.5 units of T4 DNA polymerase, in the presence of 2.5 mM dCTP and 1 μg of BSA, in a total volume of 10 μL, followed by inactivation at 75°C for 20 minutes. The product was either kept at 10°C and used within a short time, or stored at -20°C until further use.

The genes encoding the precursor peptides TrnA and TrnB from *Bacillus thuringiensis* were supplied by Genscript in a pUC57 vector. The codons were optimized for recombinant expression in *Escherichia coli*. Primers were purchased by Integrated DNA Technologies. The genes were amplified by PCR, using the pUC57 vector as template, with Takara PrimeSTAR Max DNA Polymerase (Takara Bio) (30 cycles, annealing at 55°C for 5 sec).

In 10 μL, 0.2 pmol of TrnA(B) PCR products were treated with T4 DNA polymerase following the same procedure described for pETHSUL, but substituting dCTP with 2.5 mM dGTP.

The treated inserts and vector were combined in different ratios in the smallest volume possible and incubated at room temperature for 10 minutes, before addition of 15 mM EDTA and further 5 min incubation at r.t. The mixture was directly transformed into E. cloni competent cells (Lucigen) following the manufacturer's protocol. The circular vector alone was used as a positive control and linearized and treated with T4 DNA polymerase pETHSUL as negative control.

The cells were grown in LB medium with 50 μg/mL ampicillin at 37°C for 14 hrs and the plasmid purified using Monarch Plasmid Miniprep Kit (NEB). In order to proceed with protein expression, pETHSUL-TrnA plasmid was transformed following the heat shock protocol into *E. coli* BL21 Star (DE3) cells and pETHSUL-TrnB in BL21(DE3)pLysS (no particular reason for this choice, just comparison of the expression levels).

Expression and purification of SUMO-TrnA/TrnB

BL21 cells were grown in LB medium containing 50 μg/mL ampicillin, at 37°C, 225 rpm for 12-16 hrs. They were then transferred into culture flasks and grown in 1 L LB and 50 μg/mL ampicillin at 37°C, 225 rpm, until an OD_{600} of 0.5. IPTG was added at a final concentration of 0.2 mM and the cells let express for 3 hrs at the same temperature, then harvested at 5,000 rpm for 10 min and frozen at -80°C. SUMO-TrnA(B) purification was carried out following the procedure described for TrnD, with the only difference in the P2 size-exclusion column elution buffer, being 50 mM Tris/HCl (pH 8.0) in this case. The spin-filter concentrators used were 10k MWCO. The proteins were concentrated down to ~4 mg/mL (measured by Bradford assay), because if concentrated further they started to form a white precipitate.

Expression and purification of dtUD1 protease

The plasmid pSUPER was kindly provided by Dr. P. Loll. This plasmid encodes a SUMO fusion of the dtUD1 that self-cleaves during expression in *E. coli*. Following Loll's procedure[43], pSUPER was transformed into *E. coli* BL21 Star (DE3) cells and 10 mL of LB containing 50 μg/mL ampicillin and 25 μg/mL chloramphenicol were inoculated with a single colony, and allowed to grow at 37°C, 225 rpm overnight. One milliliter of the preculture was used to inoculate

1 L of LB in culture flasks containing the same antibiotics. The cultures were grown until they reached an OD_{600} = 0.5 and IPTG was added to a final concentration on 0.2 mM. The cells were let expressing for 3-4 hrs, harvested and frozen at -80°C. The pellet was resuspended in 50 mM sodium phosphate (pH 8.0), 250 mM NaCl, 25 mM imidazole, 10% w/v glycerol, 5 mM β-mercaptoethanol and 10 mM $MgCl_2$. A protease inhibitor cocktail tablet (Thermo) was also added. The cells were lysed using ultra-sonication and the insoluble cell debris centrifuged at 20,000 rpm, 1 hr. A Ni-NTA resin column was equilibrated with the resuspension buffer. The soluble cell lysate was loaded onto the column and washed with the same buffer at 2 mL/min, until absorbance was stable at ~0.2. The protein was eluted at 5 mL/min in 50 mM sodium phosphate (pH 8.0), 250 mM NaCl, 300 mM imidazole, concentrated down with 10k MWCO centrifuge filters (Amicon) and dialyzed against 20 mM Tris/HCl (pH 8.0), 150 mM NaCl, 20% (w/v) glycerol, 1 mM EDTA and 5 mM DTT at 4°C, overnight, using SnakeSkin dialysis tubing (Thermo). The protein was diluted to a concentration of 0.25 mg/mL in the storage buffer (20 mM Tris/HCl (pH 8.0), 150 mM NaCl, 50% (w/v) glycerol, 1 mM EDTA and 5 mM DTT, divided in small aliquots and frozen at -80°C for later use. Dr. P. Loll claims that dtUD1 can be stored in these conditions at -20°C or -80°C for 6 months with no apparent loss of activity.

Cleavage of SUMO fusion protein

The SUMO fusion protein was cleaved using dtUD1 protease in a 1:200 protease/protein ratio. In order to optimize the cleavage conditions, to a 20 mL solution of ~4 mg/mL protein in 50 mM Tris/HCl (pH 8.0), NaCl, EDTA and DTT were added to a final concentration of 150 mM, 1 mM and 2 mM respectively. At r.t., 400 μg of dtUD1 were added to the protein solution and incubated for 1 hr. The solution was loaded at 5 mg/mL onto a Ni-NTA column, previously

equilibrated with the same buffer, and the flow-through was collected. When the absorbance was stable around 0.2 A, the buffer was switched to Buffer B and a second peak was collected in a separate fraction. The fractions were analyzed by SDS-PAGE and UPLC/Q-ToF MS. The supernatant was filtered through a 0.22 μm membrane and an aliquot injected onto an Agilent Zorbax C-18 column (4.6 mm X 250 mm) pre-equilibrated in 20% ACN in H_2O, 0.1% (v/v) formic acid. The peptide was eluted in a linear gradient 20-100% ACN over 8 min, 1 mL/min flow. The Q-ToF was operated in positive mode. For the MS/MS analysis, the peak corresponding to the +3 charge state was fragmented (for both TrnA and TrnB), with low collision energy.

In vitro TrnC and TrnD assays

The activity of TrnC and TrnD was tested *in vitro* using TrnA as substrate, since its purification led to a cleaner mixture compared to TrnB. Similarly to the SAM cleavage assays described, all the reactions were carried out under anaerobic conditions in an inert atmosphere glovebox. All the stock solutions were brought into the box and let incubate with nitrogen for about 1 hr. The final volume for each assay was 100 μL in 50 mM Tris/HCl, 0.1 M NaCl (pH 8.0) buffer. In order to reconstitute further the iron-sulfur clusters, 12 μM of TrnC/25 μM TrnD were incubated at r.t. for 15 min with 5 mM DTT, 200 μM ammonium iron(III) citrate and 200 μM sodium sulfide. TrnA was added to a final concentration of ~100 μM (estimated, from dtUD1 cleavage step), followed by addition of 500 μM SAM and 15 min incubation at rt. To initiate the reaction, 1 mM sodium dithionite was added and the mixture incubated at r.t. for 3-4 hrs. Controls with no SAM, or no peptide, or no reducing system were also carried out. The proteins were then denatured with 20% ethanol, the reaction tubes placed on ice for 15-20 min and the precipitate spun down in a microcentrifuge at 14,000 rpm for 10 min. The supernatant was analyzed by mass spectrometry.

Co-expression of TrnA + TrnC/D and TrnA purification

pETHSUL-TrnA was co-transformed in BL21 Star (DE3) *E. coli* cells with pET-28a(+)-TrnC and pET-28a(+)-TrnD (separately), in order to test if any modifications would occur *in vivo*. The cells were plated in LB agar plates containing 50 μg/mL ampicillin and 25 μg/mL kanamycin. The same experiment was repeated with pETHSUL-TrnB, but no colonies were observed on the plates. A colony from each plate (TrnA+C and TrnA+D) was used to inoculate two 10 mL LB cultures containing the same antibiotics. Expression was carried out as described for TrnC(D). Purification followed the procedure described for SUMO-TrnA/B, with Ni-affinity chromatography and desalting size-exclusion. As expected, TrnC precipitated as soon as exposed to oxygen, but without affecting the purpose of this study. Concentration of SUMO-TrnA was indirectly estimated: using a Bradford assay TrnD + SUMO-TrnA concentration was determined. TrnD concentration was estimated from the 410 nm absorbance of its [4Fe-4S] cluster (ε = 16,000 M^{-1} cm^{-1}) (not present in SUMO-TrnA) and subtracted to TrnD + SUMO-TrnA concentration, to obtain an estimation of SUMO-TrnA. Cleavage of the SUMO fusion was carried out following the procedure described for SUMO-TrnA(B). SUMO and TrnD were retained by Ni-resin and TrnA was collected in the flow through and concentrated. All the steps described were analyzed by SDS-PAGE. TrnA final elution was analyzed by HPLC-MS.

Results and Discussion

TrnC and TrnD

The genes *trnC* and *trnD*, which code for the presumed sactisynthases, were cloned into pET-28a(+) plasmids, in order to have a different antibiotic resistance from the two peptides, for future co-transformation experiments. They also included a N-terminal His_6-tag, with a thrombin cleavage site between the tag and the protein. Temperature and incubation time for protein expression in BL21 Star (DE3) *E. coli* were optimized for both enzymes to 5-6 hrs at 25°C, with 20 mg/L yield for TrnD and 6 mg/L for TrnC. TrnC in unstable in aerobic conditions, as reported for several radical SAM enzymes. This complication is typically attributed to protein unfolding and precipitation as a consequence of exposure to oxygen, responsible for iron oxidation and subsequent degradation of the iron-sulfur clusters[44]. TrnD did not show the same behavior and it was stable throughout purification. Nonetheless, TrnD has to be stored in relatively high concentration of thiol-reducing agents (such as β-ME) and glycerol, otherwise, the solution quickly becomes cloudy and the protein precipitates if subjected to freeze-thaw. Noticeably, TrnC and TrnD were found to be very stable over several hours at room temperature after reconstitution, if kept under strictly anaerobic conditions (inert atmosphere glovebox).

After overnight iron-sulfur cluster reconstitution in a buffer containing Fe, S and DTT, and subsequent size-exclusion chromatography to remove any excess salts, the Fe and S content was determined. The analysis resulted in 4.3 ± 0.2 equiv. of S and 5.1 ± 0.3 equiv. of Fe for TrnC, 3.2 ± 0.5 eq. S and 4.5 ± 0.2 eq. Fe for TrnD. These results are in accordance with some sactisynthases previously described and they indicate the presence of at least one [4Fe-4S] cluster for each enzyme. This cluster probably corresponds to the radical SAM Fe-S cluster (and not a possible

auxiliary one), because it is likely to be the most readily reconstituted. Moreover, the proteins would probably unfold and precipitate if the main cluster is not formed. At this stage, the presence of other Fe-S clusters cannot be excluded, because reconstitution of the clusters *in vitro* is not always possible. Auxiliary clusters might be present *in vivo* and degrade during purification. Degradation of these clusters would likely not cause the proteins to unfold, as shown in previous studies[23]. Since TrnC and TrnD sequence alignment results in several sequence identities with sactisynthases that contain auxiliary clusters, site-directed mutagenesis studies of the SAM clusters and SPASM domain cysteine residues are required to further investigate this matter. Mutation od different cysteine residues would show if the cluster which is being reconstituted is the radical SAM cluster or the auxiliary one.

A qualitative analysis of the Fe-S cluster(s) is shown in Fig. 21, TrnC and TrnD 200-700 nm absorbance profile. The characteristic absorbance trace for [4Fe-4S] cluster proteins has its maximum peak at 410 nm (ε = 16,000 M^{-1} cm^{-1})[45]. Upon addition of sodium dithionite, this peak is shown to be partially or significantly reduced for other sactisynthases[12,21], with loss of the usually brown color of the solution. TrnC and TrnD Fe-S cluster(s) were surprisingly very stable to anaerobic reduction with sodium dithionite, even after long periods of incubation (hours to

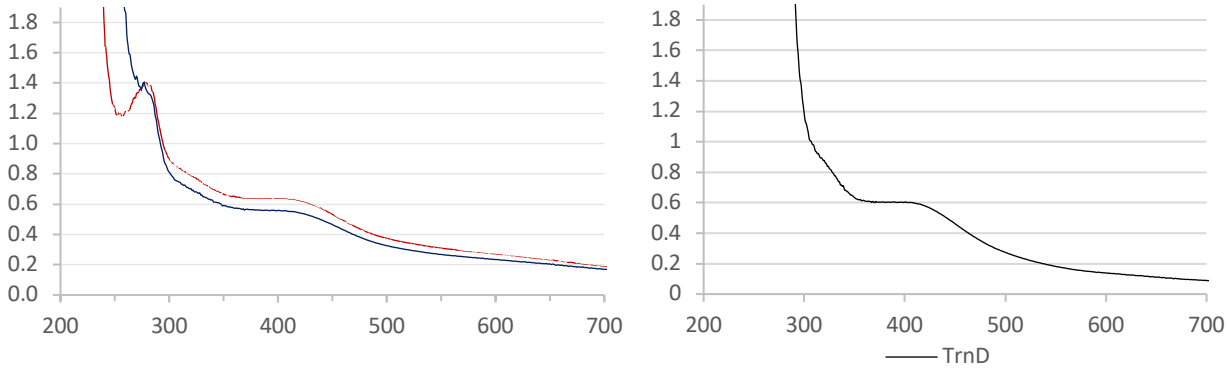


Figure 21. TrnC and TrnD UV-vis profile. In blue, TrnC Fe-S cluster(s) partial reduction by sodium dithionite. TrnD did not show any change in the spectrum upon addition of DTN.

overnight). The absorbance peak at 410 nm was reduced by only ~13% in intensity for TrnC, similar to previously reported reductions of $[4Fe\text{-}4S]^{2+}$ to $[4Fe\text{-}4S]^{1+}$ for biotin synthase[46] and YydG[36]. However, the data is not totally comparable, because an increase in absorbance at 330 nm corresponding to sodium dithionite oxidation is usually observed. In these examples, absorbance at 410 nm also increases again upon oxygen exposure. This is in contrast with the hypothesis that the SAM $[4Fe\text{-}4S]^{2+}$ cluster is the one to be reconstituted, because it would also readily reduce after addition of DTN, unless particularly stable or protected in the active site. As already mentioned, further site-directed mutagenesis studies of the cysteines would be important for a better understanding of these spectra. No significant change was observed for TrnD (not shown).

In order to understand what reducing system works best for both proteins, assays containing the protein, SAM and different reducing systems were run under anaerobic conditions, analyzed by HPLC to identify SAM background cleavage and compared to 5'-deoxyadenosine standards (Fig. 22). When incubated with sodium dithionite or titanium(III) citrate, TrnC was able to consume some of the SAM and produce detectable quantities of 5'-deoxyadenosine; no SAM cleavage was observed if TrnC was treated with the NADPH/Fld/FNR system, shown to reduce other radical SAM enzymes. TrnD did not produce 5'-dAH with any of the reducing systems. All the sactisynthases previously reported have shown the ability to cleave SAM and produce 5'-dAH upon DTN reduction, even without any substrate bound. Therefore, since TrnD also does not show any change in the Fe-S cluster UV-vis spectrum analysis when sodium dithionite is added, the SAM $[4Fe\text{-}4S]^{2+}$ cluster is either not reacting or not being reconstituted. In the second case, an auxiliary $[4Fe\text{-}4S]^{2+}$ cluster could be forming and showing in the UV-vis spectrum, but could not react with SAM.

In Fig. 23, TrnC and TrnD sequence alignment with the main members of the sactisynthase family is shown. As expected, the CXXXCXXC motif is perfectly aligned and conserved among all enzymes. Noticeably, three cysteines in the SPASM domain are 100% conserved even if at different distance from each other. As discussed previously, these Cys have been shown to coordinate an auxiliary Fe-S cluster for other sactisynthases and, since they are conserved in TrnC and TrnD, the presence of a secondary cluster is expected. TrnC and most other sactisynthases also contain a third set of three cysteines in the SPASM domain, whereas TrnD does not. TrnC shows very good sequence alignment with Tte1186, especially near the C-terminus where this enzyme has been reported to coordinate two auxiliary clusters, with strong evidences supported by mutation of several different Cys. Overall, TrnD is likely to coordinate at least one auxiliary Fe-S cluster and TrnC two. These clusters might be lost during protein purification and the proteins might not be (or only partially) reconstituting them *in vitro*. But, since TrnC was shown to cleave SAM and form 5'-dAH under reducing conditions, the $[4Fe\text{-}4S]^{2+}$ cluster coordinated to the CXXXCXXC motif is likely to be undergoing reduction, despite the ambiguity of the UV-vis spectral data.

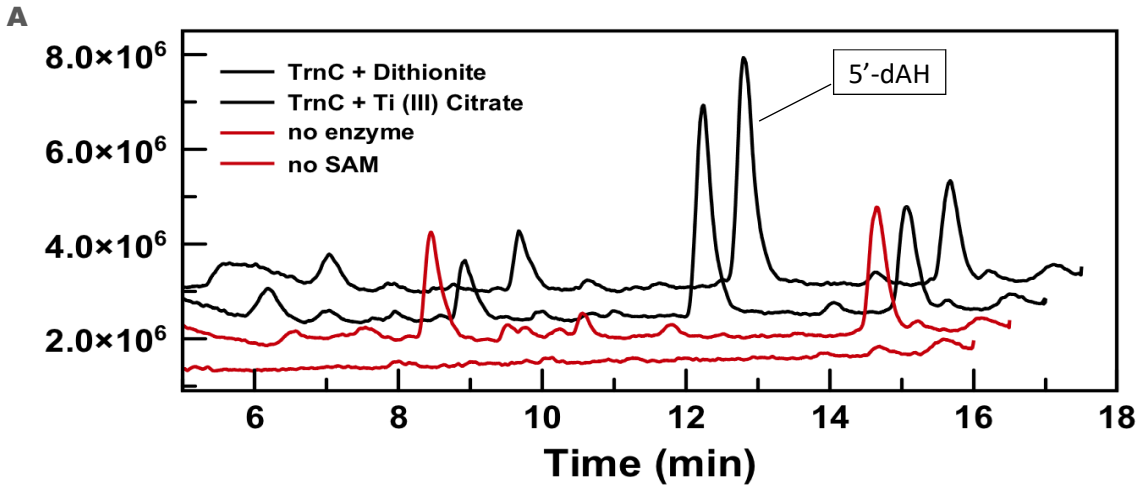


Figure 22. A. SAM cleavage assay. Total Ion Chromatogram (TIC) of SAM cleavage assays with TrnC; 5'-deoxyadenosine elutes at 11.0 min. In red, controls where either TrnC or S-adenosylmethionine were not added. For the controls, no 5'-dAH is observed.

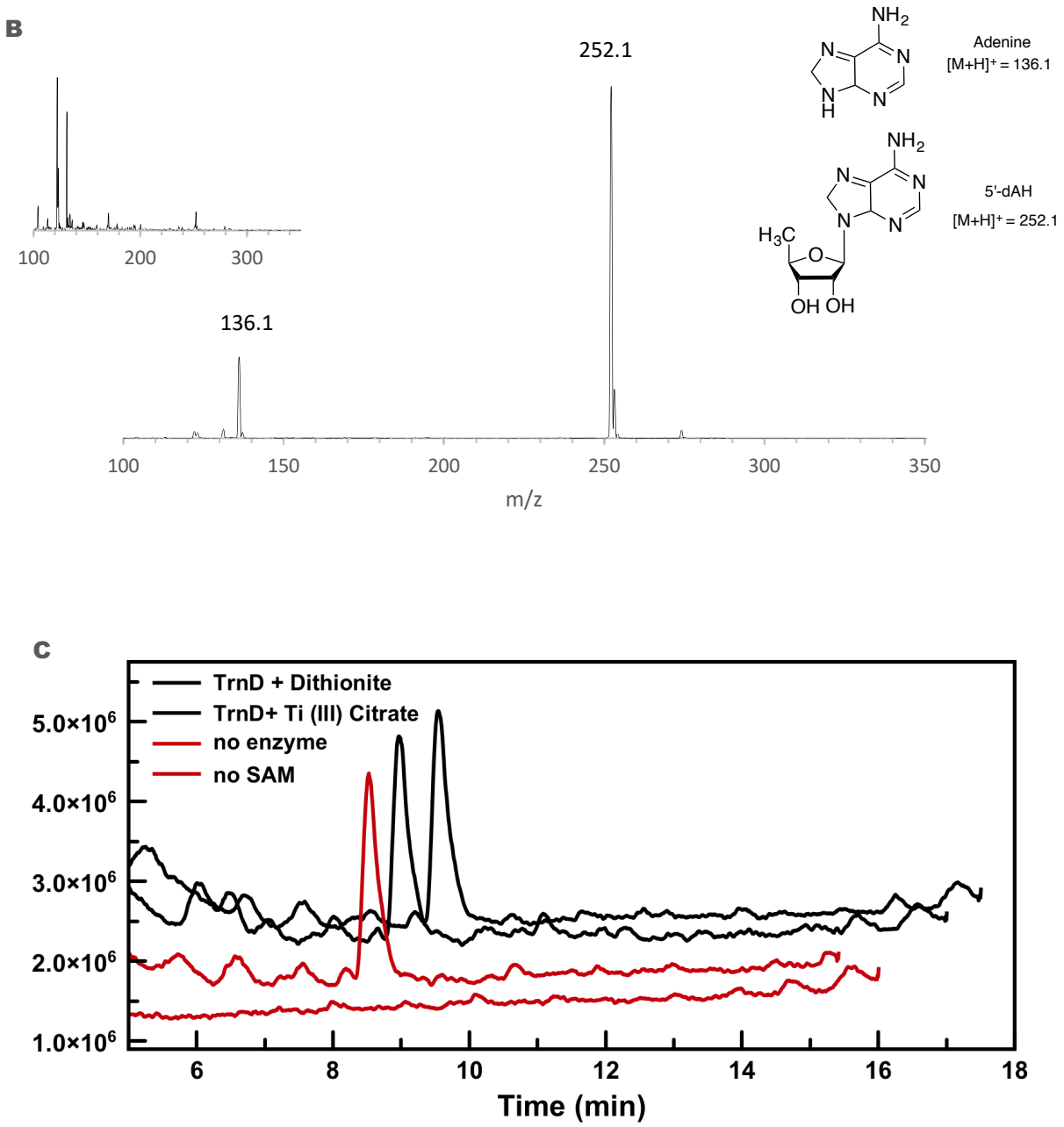


Figure 22. B. 5'-deoxyadenosine extracted-ion chromatogram. $[M+H]^+ = 252.1$ is the major peak and fragmentation to adenine (136.1) the minor peak, as expected. 253.1, corresponding to 5'-dAH ^{13}C isotopes is also detected. **C.** TIC of SAM cleavage assay with TrnD. No 5'-dAH is detected with any of the reducing systems used.

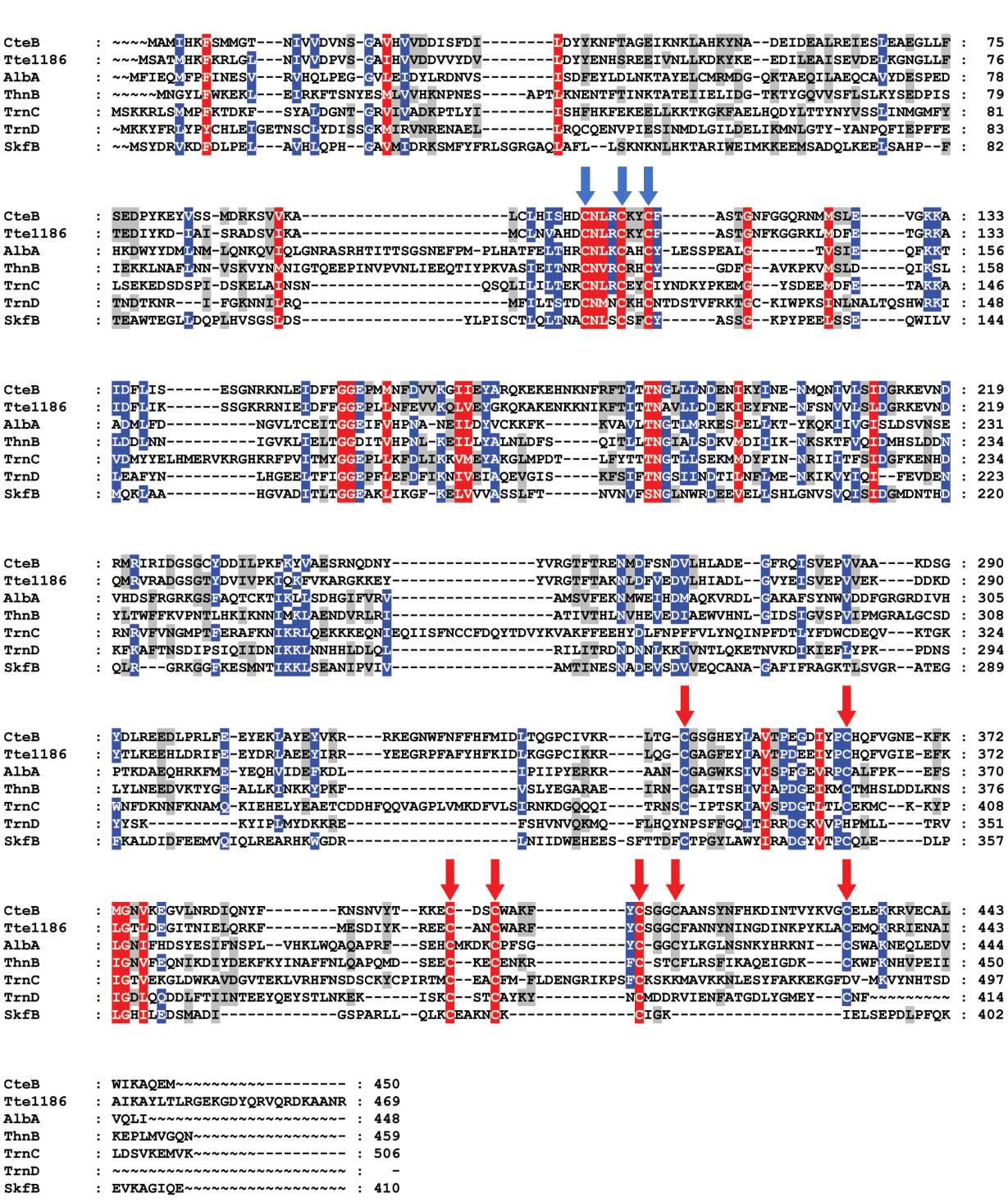


Figure 23. TrnC and TrnD MUSCLE sequence alignment. MUSCLE sequence alignment including TrnC, TrnD and most of the sactisynthases described. The blue arrows indicate the radical SAM domain, whereas the red arrows indicate cysteine residues in the SPASM domain, possibly involved in coordinating one or two auxiliary clusters. Red = 100% conserved, blue = 75%, gray = 50%.

TrnA and TrnB

Expression and purification of TrnA and TrnB was very challenging due to their poor solubility in water and small size. Both peptides are comprised of a majority of hydrophobic residues, with only a few charged amino acids concentrated near the N-terminus. The natural product thuricin CD reported by Rea *et al.*[27] was isolated from *B. thuringiensis* cell pellet by extraction in 70% aqueous isopropanol (acidified to pH 2.0), relatively harsh conditions which indicate an issue in solubility. The first attempt in expression and purification of thuricin CD precursors was to clone TrnA(B) into a pET plasmid, and transform and express in *E. coli*. Following the procedure just described, the cell pellets were extracted in 70% isopropanol (pH 2.0), but none of the peptides were detected by further analyses. The reasons for this are probably the following: Thuricin CD is likely to be more soluble than its linear precursors TrnA and TrnB, due to its more compact partial secondary structure imparted by the thioether crosslinks and shorter length (~20 residues shorter); expression in *E. coli* might not be optimal compared to *B. thuringiensis*; lastly, the linear precursors are more likely to be subjected to protease degradation during expression. In regard to Subtilosin A, Tte1186a and CteA, the precursor peptides were simply His-tagged and purified by Ni-affinity chromatography in aqueous conditions. However, addition of a N-terminal His_6-tag to TrnA and TrnB and purification through Ni-affinity chromatography, did not result in any detectable peptide.

MBP-TrnB

The first positive results in TrnA(B) isolation were found with the introduction of a fusion protein at the N-terminus of the peptide. Initial attempts were focused on TrnB, because its amino acid sequence contains a tryptophan residue, making this peptide more UV-active compared to

TrnA, which does not contain any residue with a chromophore. The plasmid selected was pMAL-c2x, expressing a maltose-binding protein (MBP), which may also improve the purification step, since it requires the use of amylose-affinity chromatography, shown to result in a lower non-specific protein background binding in some cases, compared to Ni-affinity[47]. Using this type of resin also avoids high imidazole concentrations in the elution buffer, which would require subsequent purification steps. Low concentrations of maltose, compatible with further experiments, are sufficient to elute the protein. Therefore, TrnB was cloned into pMAL-c2x at the C-terminus of MBP. From the commercial vector, the cleavage sequence between MBP and TrnB was mutated to a Tobacco Etch Virus (TEV) protease cleavage site, with the addition of an extra serine at the N-terminus of the target peptide. MBP and TEV protease were both His_6-tagged at the N-terminus, in order to be purified together after the final cleavage step. As shown in the SDS-PAGE in Fig. 24, the protein construct was successfully purified, but during expression, or lysis and centrifugation, a large fraction of MBP-TrnB (~47,000 kDa) was degraded to MBP only (~42,000 kDa) (lane 1). In fact, after TEV protease cleavage (lane 2), the MBP-TrnB band combines to the MBP band. An unknown contaminant below the 40 kDa marker is also detected. Proteolysis degradation is a common issue for small target proteins/peptides expressed with a fusion partner. TrnB is easily degraded in this case, because it is a short linear peptide protruding from a large globular protein with well-defined structure. Different expression time and temperatures did not make any significant difference. From the gel, the degradation issue seems to decrease the final purification yield approximately by half. Moreover, TrnB is only 10% by mass

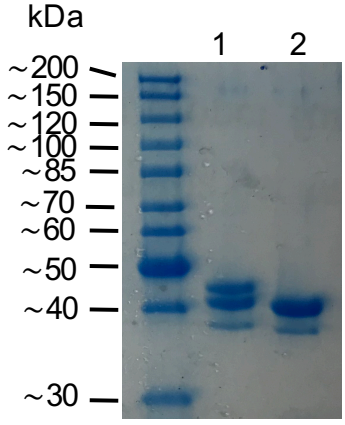


Figure 24. MBP-TrnB SDS-PAGE. Lane 1: full length MBP-TrnB; lane 2: after TEV protease cleavage.

of the total fusion MBP-TrnB protein. Therefore, for example, with ~100 mg/L yield after purification of MBP-TrnB, and half corresponding to the truncated protein, total TrnB after TEV protease cleavage would be only ~5 mg/L. TEV protease activity is also not always optimal with such globular proteins, and it requires further Ni-affinity chromatography purification, decreasing the final yield even more. The amount of TrnB isolated was simply too low to carry out further experiments and a different strategy was adopted.

MBP-TrnB-CBD

In order to prevent proteolytic degradation, a second double fusion protein system was engineered, with the idea of having the peptide between two well expressing proteins. For this purpose, an intein-tag containing a chitin-binding domain (CBD) was attached to TrnB C-terminus. This type of fusion protein is easily cleavable upon addition of high DTT concentrations (40 mM or higher) and can be purified by chitin-affinity chromatography. Although this choice adds a second cleavage and purification step, degradation of TrnB from the C-terminus would be completely avoided (Fig. 25). The manufacturer suggests running the CBD cleavage on a chitin-resin column, but yields were much higher for cleavage in solution for our construct. However, after the final TEV protease cleavage, the peptide yield was still very low, and it was not possible to recover TrnB from the final Nickel

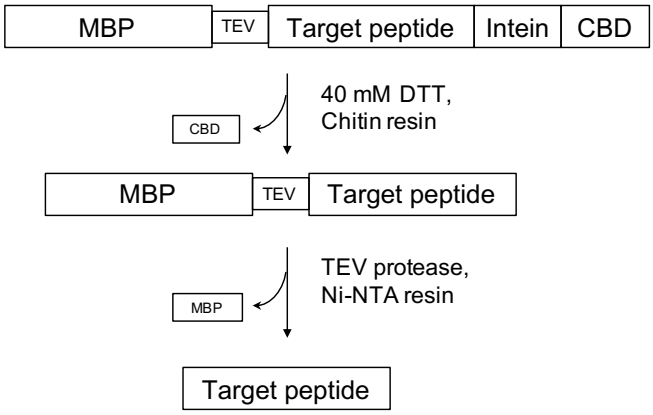
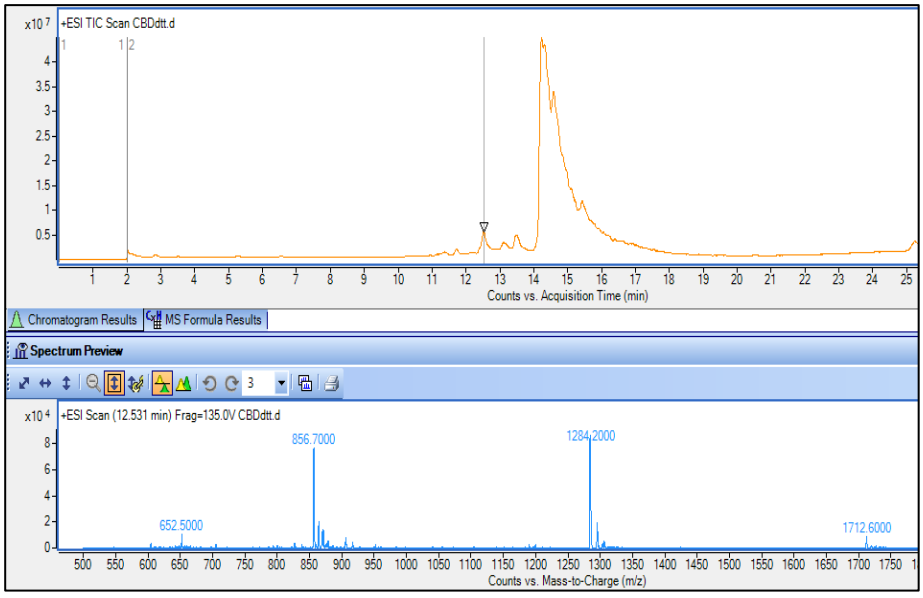


Figure 25. MBP-TrnB-CBD purification scheme. TrnB is eluted in the final Ni-column flow-through. "TEV" in the figure indicates TEV protease recognition sequence.

resin purification step. As reported for other short hydrophobic peptides expressed with a fusion partner, HPLC purification was attempted. The UV trace resulted in several peaks that were collected separately, concentrated and analyzed by mass spectrometry. The peptide was identified, but the thiolates of all three cysteines required for activity were irreversibly oxidized to sulfonic acid, making TrnB soluble in water, but the thioether crosslink formation reaction impossible to study. The peaks shown in Fig. 26 correspond to the $[M+3]^{3+}$ (1712.60), $[M+4]^{4+}$ (1284.20) and $[M+6]^{6+}$ (856.70) charge states of TrnB plus nine oxygen atoms, three for each cysteine. This oxidation probably occurred during the overnight cleavage of CBD step with DTT, during which oxygen oxidized the cysteines, and such high DTT concentration, probably containing some oxidized DTT, also acted as a catalyst for cysteine oxidation. A new approach that avoids high concentrations of DTT was needed, and CBD as fusion protein was not further considered a viable option.



TrnB, all Cys oxidized: $[M+3]^{3+}$ (1712.60), $[M+4]^{4+}$ (1284.20), $[M+6]^{+6}$ (856.70)

Figure 26. TrnB Cys oxidation.

MBP-KR5/10-TrnB

Berteau *et al*. successfully purified a truncated version of polytheonamide A at the N-terminus[48], a peptide with a short sequence of hydrophobic residues, after the addition of five lysine residues at the N-terminus. Applying a similar strategy, TrnB solubility was increased by the addition of five (MBP-KR5-TrnB) and ten (MBP-KR10-TrnB) cationic residues (a combination of lysine and arginine) at the amino end. Since all TrnC/D modifications form near the C-terminus, addition of these residues was thought to only minimally affect TrnB reactivity. Following the procedure described for MBP-TrnB, the expression levels were drastically low and recovery of only ≤ 1 mg/L of the fusion protein was possible, still contaminated with the truncated version (MBP). Again, after the TEV cleavage, it was impossible to isolate the peptide with Ni-affinity chromatography. TrnB was also not detectable by MS from the crude cleavage mixture,

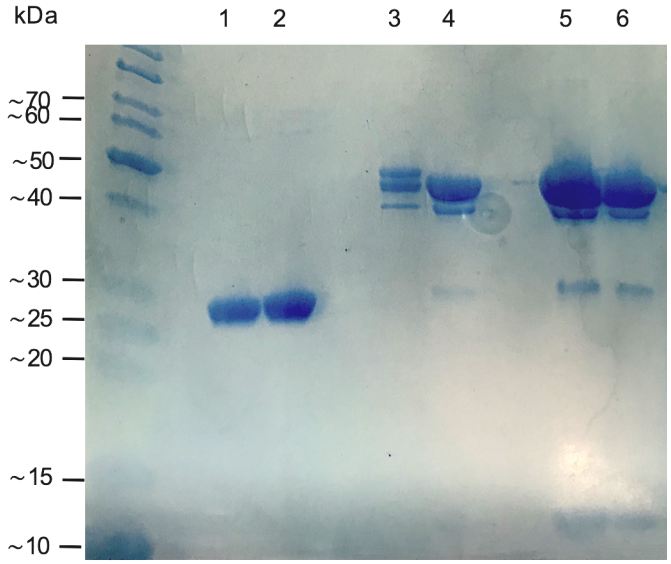


Figure 27. SDS-PAGE, MBP-KR5/10-TrnB. Lane 1 and 2: TEV protease; lane 3: purified MBP-KR5/10-TrnB; lane 4, 5, 6: different amounts of MBP-KR5/10-TrnB after TEV protease cleavage. The peptide is the faint band slightly above the 10 kDa band of the ladder.

probably due to its little amount compared to MBP and TEV proteins, and hard to separate under their broad peaks. However, for the first time it was possible to reproducibly identify the peptide by 8-16% acrylamide gradient SDS-PAGE, using a modified Coomassie Blue gel dye solution

(AquaStain, Bulldog Bio) (Fig. 27). Due to the inaccuracy of SDS-PAGE in analyzing small molecules, the peptide runs with the smallest band of the ladder (~10 kDa) and seems to be larger than it should. Aprotinin (~6kDa) was used as a standard and KR5/KR10-TrnB (~5.5 kDa) was perfectly comparable to it. Unfortunately, due to the very low expression and purification yields, it was not possible to carry out further experiments with these compounds, and a new strategy was adopted. Additionally, the number of positive charges, be it five or ten Arg/Lys residues, did not make a significant difference overall.

SUMO-TrnA/TrnB

A different approach was taken, exchanging the N-terminal fusion partner MBP for a smaller protein with more optimal final protease cleavage and purification steps, following similar successful sactipeptides purification approaches reported[18, 21]. Marahiel *et al.* use thioredoxin (TXN) as a fusion protein, which is approximately 11.7 kDa and a TEV or HRV-3C protease cleavage site. TEV and HRV-3C perform sequence specific proteolysis and require subsequent HPLC purification of the peptide. Dr. P. Loll kindly donated to us plasmids encoding Small Ubiquitin-like Modifier (SUMO) fusion protein, a small protein (~12 kDa) that can be cleaved by a highly specific protease, which recognizes the tertiary structure of SUMO rather than

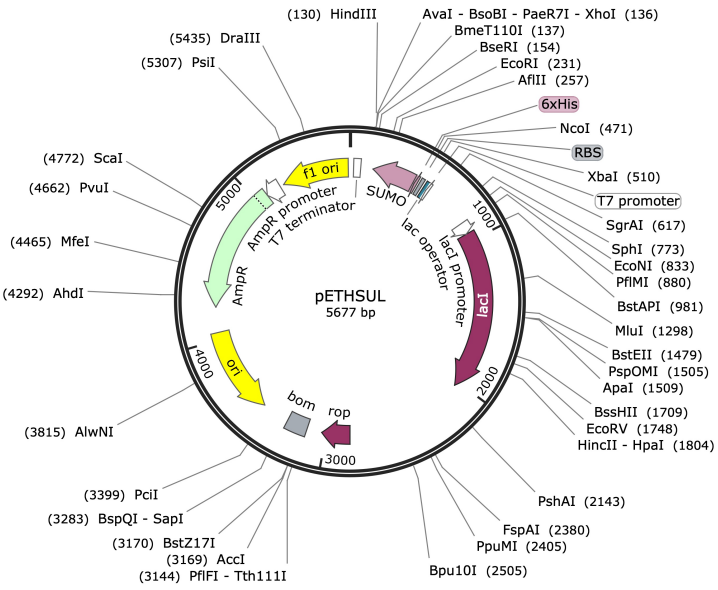


Figure 28. pETHSUL vector. The genes *trnA* and *trnB* were inserted in frame with the SUMO fusion protein, using LIC (BseRI restriction digest site).

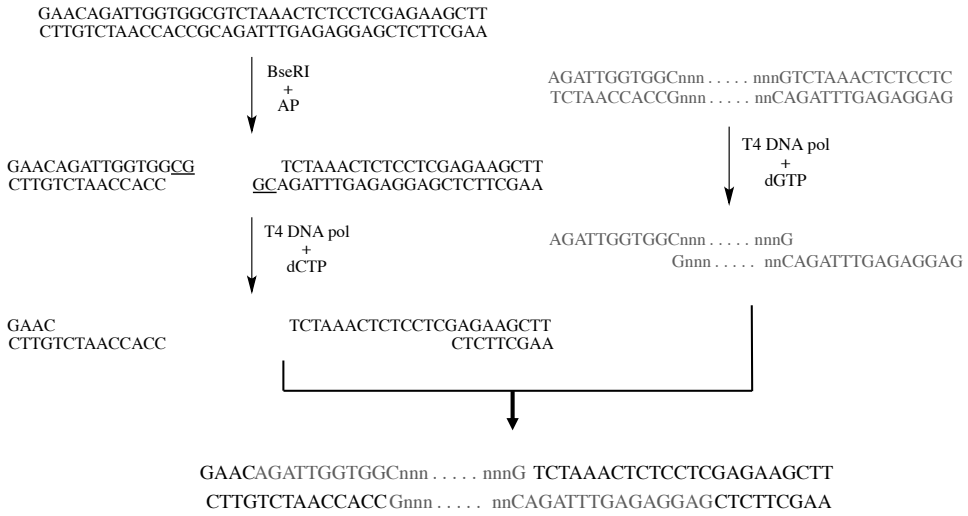


Figure 29. Ligation independent cloning workflow for pETHSUL-TrnA(B).

an amino acid sequence[49]. The plasmid encoding this protein was also donated to us, and both SUMO and dtUD1 (dtUD1, doubly-tagged ubiquitin domain 1 protease) contained a His_6-tag, convenient for purification. The *E. coli* codon-optimized genes *trnA* and *trnB* were cloned into the vector through ligation independent cloning (LIC), at the SUMO C-terminus, following a modified version of the procedure published by the Loll group[43]. The plasmid selected was pETHSUL (Fig. 28), where pET is the parent vector, H stands for hexa-histidine tag and SUL for SUMO LIC. The LIC process (outlined in Fig. 29) was chosen to avoid restriction enzyme/T4 ligase traditional methods, which are time consuming and often only few genes are successfully inserted, requiring

pETHSUL-TrnA	F-primer	5'- **AGATTGGTGGC**ATGGAAGTCATGAACAATGCACTGATTACC -3'
	R-primer	5'- **GAGGAGAGTTTAGAC**TTAACCCAGCGTAAACGCGGTG -3'
pETHSUL-TrnB	F-primer	5'- **AGATTGGTGGC**ATGGAAGTCCTGAACAAACAGAACGTG -3'
	R-primer	5'- **GAGGAGAGTTTAGAC**TTACAGGAAATAAGATGCCGCCGCAA -3'

Table 1. pETHSUL-TrnA(B) cloning primers.

extra steps, such as colony PCR. The vector is linearized using the restriction enzyme BseRI, then treated with T4 DNA polymerase, which, in the presence of dCTP only, generates 10- and 14-base single strand overhangs through its 3' → 5' exonuclease activity. The target gene is amplified using primers containing a region complementary to this overhang and the nucleotide sequence required for annealing to the template. The primers used are shown in Table 1. The reverse primers also include a termination sequence, because the final peptide would contain an extra valine residue at the C-terminus otherwise. PCR products were treated with T4 DNA polymerase in the presence of dGTP, to generate the complementary single stranded overhangs. The treated vector and insert were mixed together and, after short incubation, simply transformed into *E. coli*. competent cells, that naturally repair the four resulting nicks. Competent cells prepared with calcium chloride from BL21 or DH5α cells grown from glycerol stocks would not always result in successful transformation. Therefore, commercial competent cell strains with higher transformation efficiency, such as Stellar (Takara) or Ecloni 10G (Lucigen) were utilized instead. Even with these strains, the transformation yielded only one to ten colonies per LB agar plate. pETHSUL restriction digest with BseRI was not complete (Fig. 30), therefore, the linearized vector (lane 2, 5677 bp)

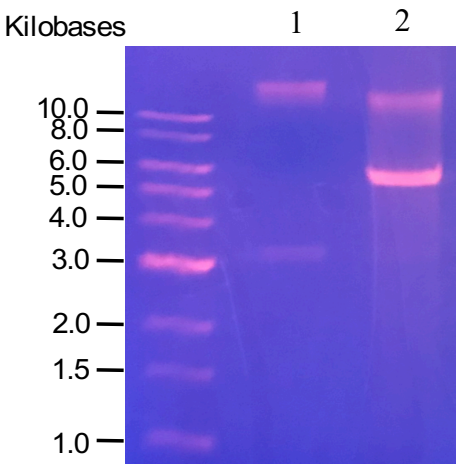


Figure 30. pETHSUL BseRI restriction digest.
Lane 1: undigested vector.
Lane 2: BseRI digest.

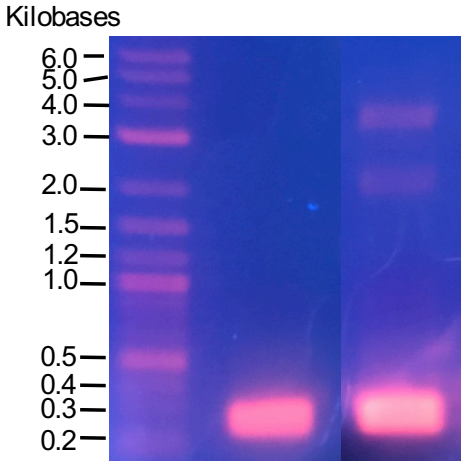


Figure 31. TrnA and TrnB PCR products. Generated using primers for LIC. TrnB band (right lane, ~0.2 kb) was purified from the agarose.

was purified using an agarose gel purification kit (GE Healthcare). This step was necessary because transformation would be much more selective for the contaminating circular plasmid, compared to the new construct formed by LIC. TrnA and TrnB PCR products from the pUC57 plasmid, using the primers indicated in Table 1 are shown in Fig. 31. TrnB (on the right) was gel purified as well. DNA sequencing (using the same set of primers) confirmed insertion of the genes. Using this construct, the SUMO-TrnA/TrnB protein expression levels increased to ~120 mg/L, measured by Bradford protein assay. Noticeably, the constructs are slow in fully reacting with the Bradford dye (4-5 minutes). Since the SUMO protein is much smaller than MBP, the peptide corresponds to ~30% of the total mass, underscoring the increased yield in this system compared to the MBP-TrnB construct. The SUMO protease dtUD1 plasmid was transformed into BL21 Star (DE3) *E. coli* cells, expressed and purified following Loll *et al.* procedure, and finally stored in low concentrations (0.25 mg/L), in a high DTT (10 mM) and glycerol (20%) content buffer. A ratio of 1:200 dtUD1/SUMO-TrnA(B) was sufficient to completely cleave the SUMO fusion protein in one hour, at room temperature. Overall, this step was also an improvement in time consumption and cleavage efficiency in particular. The cleavage mixture was loaded onto a Ni-affinity resin column by FPLC, and collection of the flow-through after dtUD1 cleavage afforded a broad absorbance peak for both peptides. Noticeably, TrnA(B) elution fraction resulted in a pale-yellow solution, perhaps due to unspecific Ni binding, but the reasons for this color have not being explored further. TrnA and TrnB were identified by SDS-PAGE (Fig. 32), showing a band shift corresponding to the loss of ~5,000 Da between the SUMO construct before and after cleavage. The Ni-resin elution band, after dtUD1 cleavage, also run as far as the SUMO-only control, purified after expression of the pETHSUL vector. PAGE shows how some proteolysis during

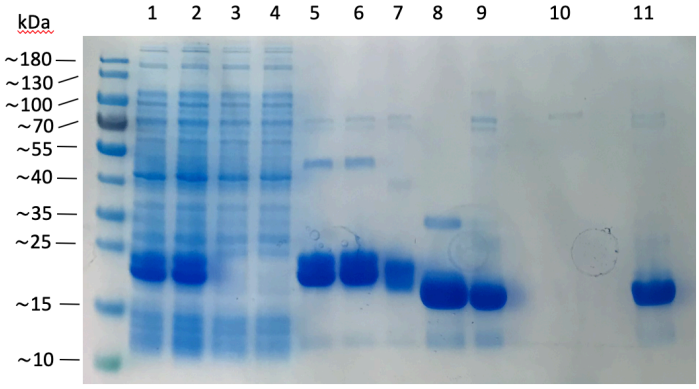


Figure 32. SDS-PAGE, SUMO-TrnA(B).

Lane 1 SUMO-TrnA lysate;
lane 2 SUMO-TrnB lysate
lane 3, 4: Ni-resin flow-through;
lane 5, 6, 7: Ni-resin elution;
lane 8: SUMO no fusion (control);
lane 9: crude dtUD1 cleavage mixture;
lane 10: last Ni-resin flow-through;
lane 11: last Ni-resin elution (+ imidazole).

expression might be occurring with this construct, as the SUMO-TrnA(B) bands are broad and seem to be almost two separate bands. But the reference SUMO-only band is also very broad, therefore it could just be a matter of amount of protein loaded.

TrnA and TrnB MS analysis

The Ni-resin flow-through elution fractions containing TrnA and TrnB after final dtUD1 cleavage were analyzed by HPLC-MS. The spectrum for each peptide resulted in a complex mixture of several truncated peptides, losing fragments of different lengths from the C-terminus. The main peaks are summarized in Fig. 33 (TrnA) and Fig. 34 (TrnB). The full-length peptides are mainly present as $[M+3H]^{3+}$, but they are in considerably low amount compared to the truncated peptides. MS/MS analysis was carried out targeting the $[M+3H]^{3+}$ for fragmentation. Most of the b and y fragments were successfully recognized, as shown in Fig. 35 (TrnA) and Fig. 36 (TrnB) and summarized in Table 2 and 3 respectively. The reasons for degradation of the full-length peptide have not been explored thoroughly, but from the elongated shape of the bands on the polyacrylamide gel, the assumption can be made was that this degradation occurs via endogenous proteases during protein expression or cell lysis. This phenomenon is still observed at different time and temperature variation during expression, and the addition of different protease

inhibitor does not seem to have any effect. The hypothesis that SUMO protease degrades the peptides by non-specific cleavage was considered, but since this protease shows >85% specificity in cleaving at the +1 position of SUMO C-terminus residue (unless it is a proline, which is not the case)[49,50], it seems unlikely. Another hypothesis is spontaneous degradation. The peptides do not encounter any particularly harsh condition during the purification process, therefore this option does not seem valuable. Noticeably, the most abundant peak is for both TrnA and TrnB a truncation at the amino acid residue in the position right before the first cysteine.

Fragment	Time (min)	Ion	Predicted	Observed	Diff.
MEVMNNA	1.284	$[M+H]^{+}$	808.33283	808.3450	0.01217
MEVMNNALIT	1.690	$[M+H]^{+}$	1135.54864	1135.5518	0.00316
		$[M+H]^{2+}$	568.27798	568.2797	0.00172
MEVMNNALITKVDEEIGGNAA	1.813	$[M+H]^{2+}$	1110.0378	1110.0429	0.0051
		$[M+H]^{3+}$	740.36098	740.3637	0.00272
MEVMNNALITKVDEEIGGNAACVIGCIGSCVI	2.111	$[M+H]^{2+}$	1633.7836	1633.8143	0.0307
		$[M+H]^{3+}$	1089.5248	1089.5466	0.0218
		$[M+H]^{4+}$	817.39545	817.4186	0.02315
Full-lenght TrnA	2.376	$[M+H]^{3+}$	1552.7628	1552.7979	0.0351
MEVMNNALITKVDEEIGGNAACVIGCIGSCVISEGIGSLVGTAFTLG		$[M+H]^{4+}$	1164.8239	1164.8492	0.0253

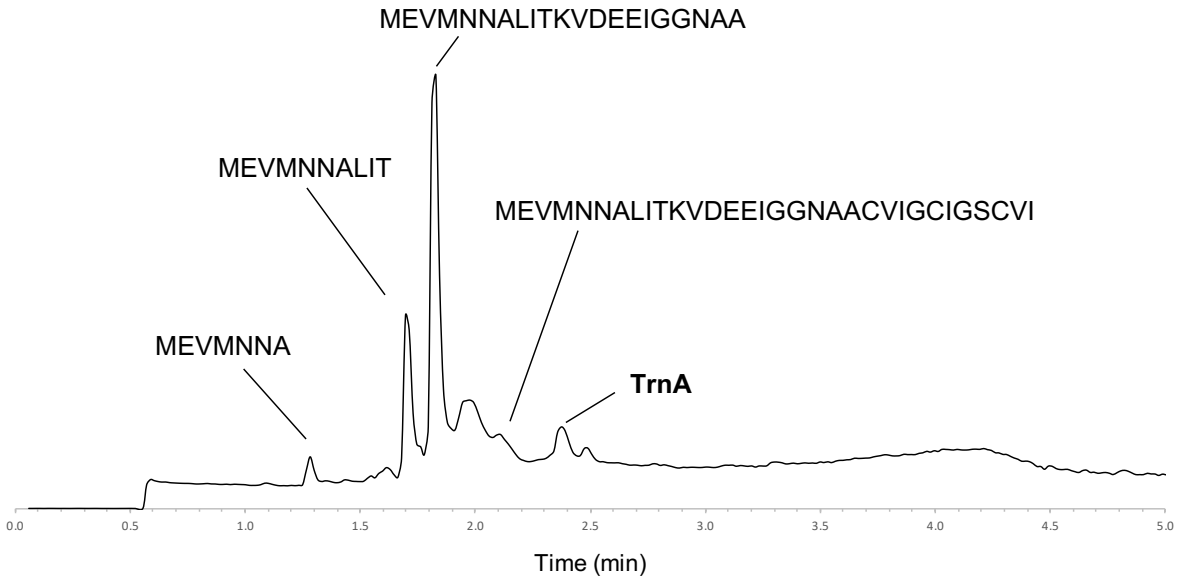


Figure 33. TrnA and TrnA degradation fragments identified by mass spectrometry.

Fragment	Time (min)	Ion	Predicted	Observed	Diff.
MEVLNKQNVNIIPESEEVGGWVA	1.836	$[M+H]^{2+}$	1278.14419	1278.1410	-0.00319
		$[M+H]^{3+}$	852.4319	852.4300	-0.0019
		$[M+H]^{4+}$	639.57575	639.5762	0.00045
MEVLNKQNVNIIPESEEVGGWVACVG	1.968	$[M+H]^{2+}$	1407.69327	1407.6928	-0.00047
		$[M+H]^{3+}$	938.79825	938.8008	0.00255
		$[M+H]^{4+}$	704.35052	704.3553	0.00478
MEVLNKQNVNIIPESEEVGGWVACVGACGTVCL	2.232	$[M+H]^{2+}$	1731.33227	1731.3307	-0.00157
		$[M+H]^{3+}$	1154.55729	1154.5586	0.00131
		$[M+H]^{4+}$	866.1698	866.1711	0.0013
Full-lenght TrnB	2.332	$[M+H]^{3+}$	1664.1356	1664.1266	-0.009
MEVLNKQNVNIIPESEEVGGWVACVGACGTVCLASGGVGTEFAAASYFL		$[M+H]^{4+}$	1248.34898	1248.3542	0.00522

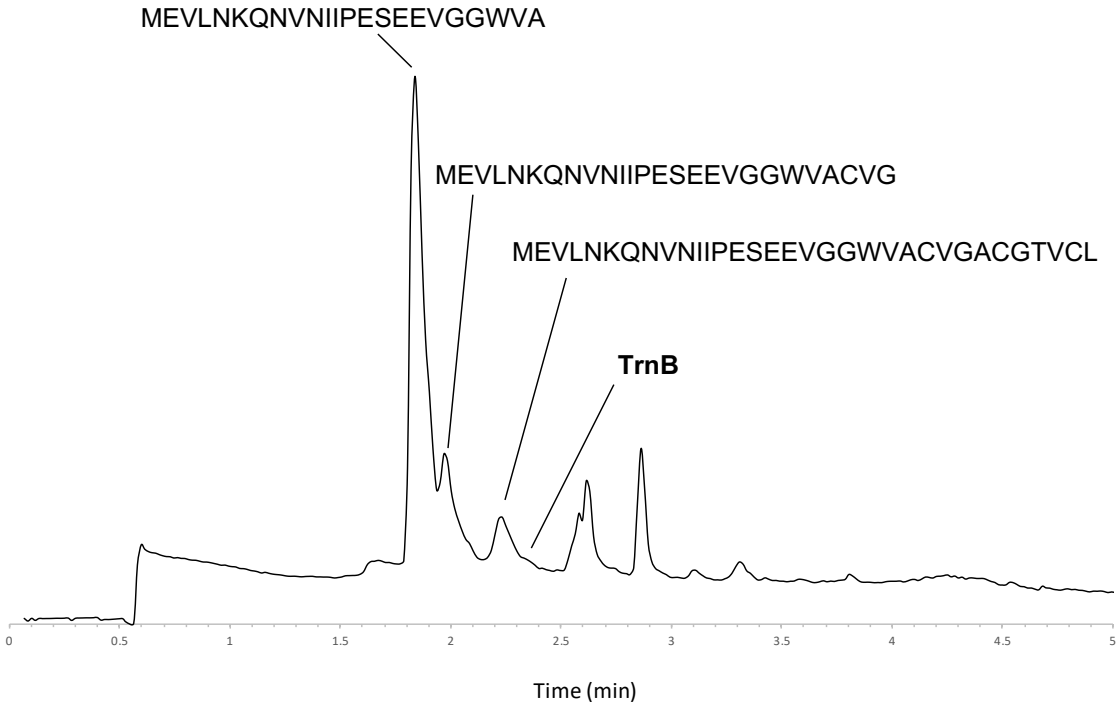


Figure 34. TrnB and TrnB degradation fragments identified by mass spectrometry.

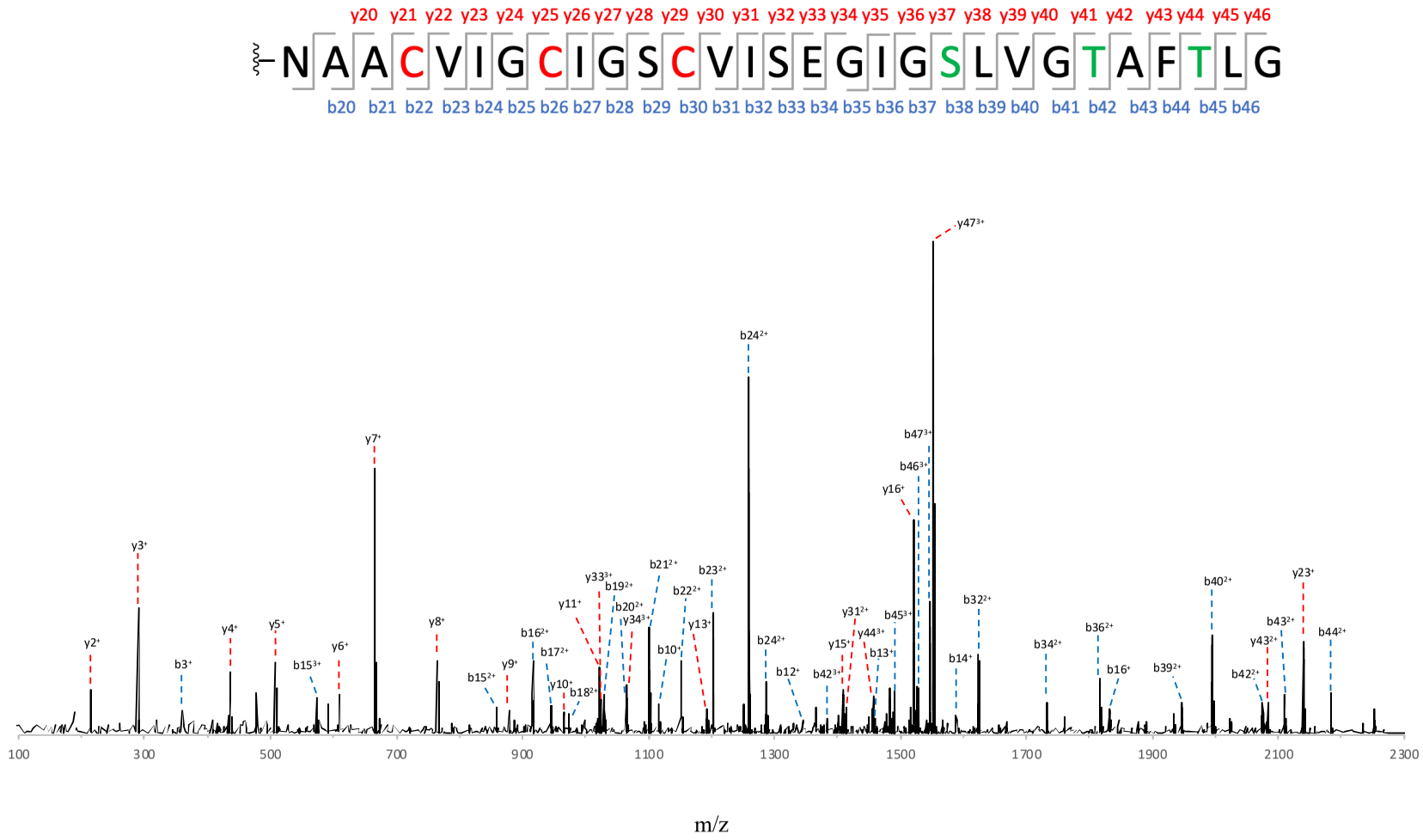


Figure 35. TrnA MS/MS fragmentation spectrum.

Fragment	Predicted	Observed	Difference	Intensity
$y2^{+}$	189.1234	189.1222	-0.0012	605.3951
$y3^{+}$	290.1711	290.1694	-0.0017	3428.9375
$b3^{+}$	360.1588	360.1595	0.0007	637.5079
$y4^{+}$	437.2395	437.2416	0.0021	1704.0376
$y5^{+}$	508.2766	508.2756	-0.0010	1933.2104
$b15^{3+}$	573.2761	573.3014	0.0253	1013.2607
$y6^{+}$	609.3243	609.3291	0.0048	1094.2686
$y7^{+}$	666.3458	666.3425	-0.0033	7156.6802
$y8^{+}$	765.4142	765.4129	-0.0013	2013.0629
$b15^{2+}$	859.4105	859.4025	-0.0079	736.5558
$y9^{+}$	878.4998	878.4903	-0.0095	677.8163
$b16^{2+}$	915.9525	915.9524	-0.0001	1149.4702
$b17^{2+}$	944.4632	944.4578	-0.0054	769.7198
$y10^{+}$	965.5303	965.5259	-0.0044	595.8893
$b18^{2+}$	972.9739	972.9651	-0.0088	572.0882
$y11^{+}$	1022.5517	1022.5491	-0.0026	1814.2703
$y33^{3+}$	1023.5083	1023.5552	0.0469	963.5298
$b19^{2+}$	1029.9954	1029.9893	-0.0061	1101.2577
$b20^{2+}$	1065.5140	1065.5057	-0.0083	954.2512
$y34^{3+}$	1066.5225	1066.5121	-0.0104	986.3112
$b21^{2+}$	1101.0325	1101.0349	0.0024	1736.5237
$b10^{+}$	1117.5381	1117.5321	-0.0060	830.4031
$b22^{2+}$	1152.5371	1152.5280	-0.0091	502.0988
$y13^{+}$	1192.6573	1192.6441	-0.0132	454.7336
$b23^{2+}$	1202.0713	1202.0630	-0.0083	2595.2898
$b24^{2+}$	1258.6134	1258.6098	-0.0036	7149.2104
$b25^{2+}$	1287.1241	1287.1248	0.0007	1413.6801
$b12^{+}$	1344.7015	1344.6930	-0.0085	417.5200
$b42^{3+}$	1383.6731	1383.6638	-0.0093	458.7234
$y15^{+}$	1408.7319	1408.7285	-0.0034	1220.5385
$y31^{2+}$	1413.6954	1413.6732	-0.0222	757.7804
$b44^{3+}$	1456.3749	1456.3582	-0.0167	1045.6371
$b13^{+}$	1459.7289	1459.7061	-0.0228	440.2130
$b45^{3+}$	1490.0575	1490.0510	-0.0065	1166.5560
$y16^{+}$	1521.8159	1521.7410	-0.0749	5347.0176
$b46^{3+}$	1527.7521	1527.7555	0.0034	1265.0002
$b47^{3+}$	1546.7593	1546.7506	-0.0087	3598.2422
$y47^{3+}$	1552.7628	1552.7513	-0.0115	12072.1572
$b14^{+}$	1588.7710	1588.7504	-0.0206	535.0822

Table 2. List of TrnA MS/MS fragments.

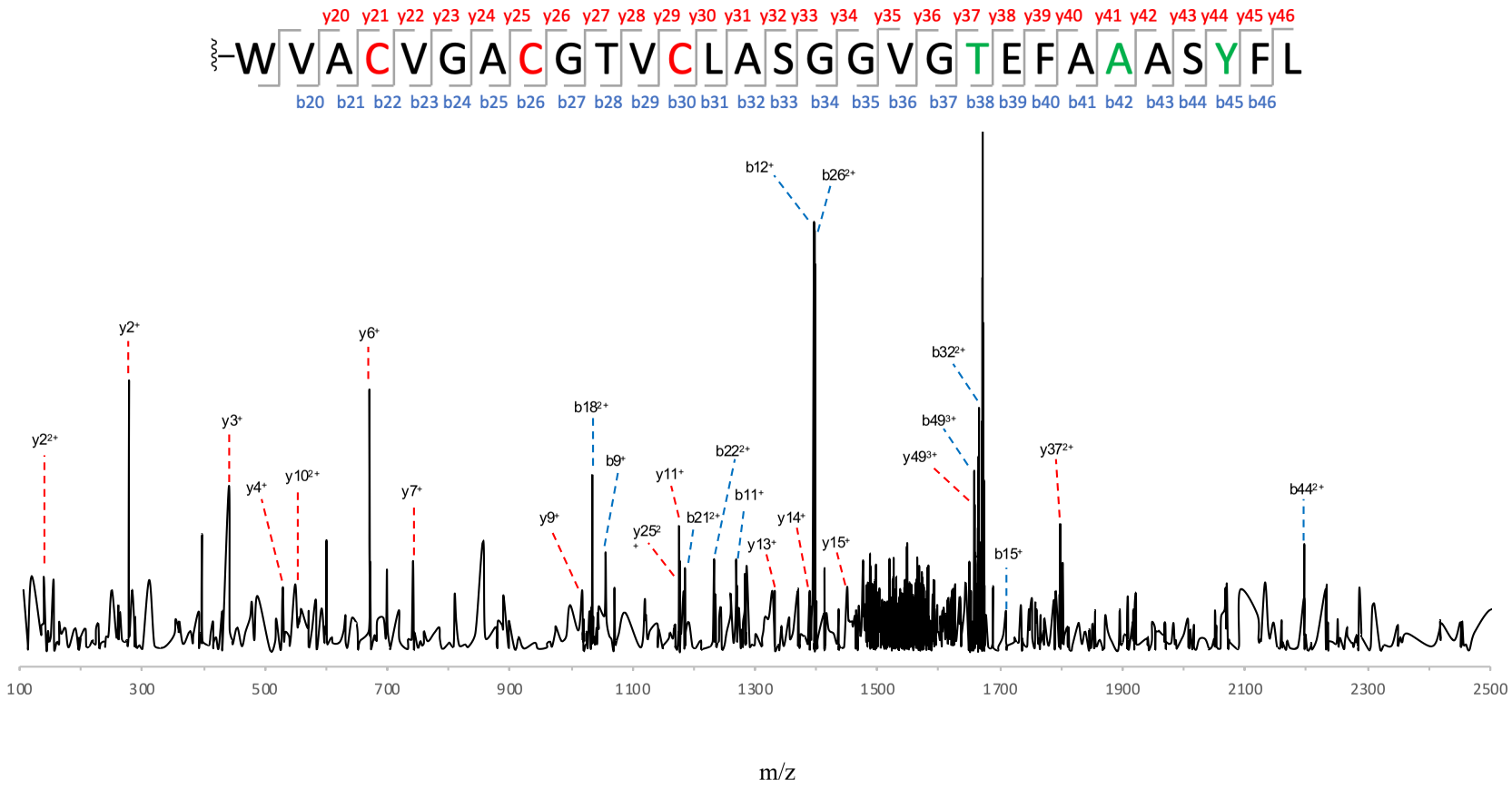


Figure 36. TrnB MS/MS fragmentation spectrum.

Fragment	Predicted	Observed	Difference	Intensity
$y2^{2+}$	140.08884	140.0669	-0.0219	300.1608
$y2^{+}$	279.17036	279.174	0.0036	969.529053
$y3^{+}$	442.23369	442.2307	-0.0030	610.377136
$y4^{+}$	529.25572	529.2584	0.0027	226.697586
$y10^{2+}$	560.27153	560.2259	-0.0456	191.580307
$y6^{+}$	671.33995	671.3398	-0.0002	933.504822
$y7^{+}$	742.37706	742.3684	-0.0087	352.339111
$y9^{+}$	1018.48807	1018.4734	-0.0147	255.276672
$b18^{2+}$	1034.02502	1034.0113	-0.0137	202.335251
$b9^{+}$	1056.55068	1056.5449	-0.0058	384.140747
$y25^{2+}$	1176.056	1176.0806	0.0246	197.235397
$y11^{+}$	1176.55721	1176.563	0.0058	475.125
$b21^{2+}$	1184.08614	1184.0779	-0.0082	251.568237
$b22^{2+}$	1233.62035	1233.5951	-0.0252	359.301178
$b11^{+}$	1283.67768	1283.6706	-0.0071	309.139526
$y13^{+}$	1332.64709	1332.6421	-0.0050	253.495377
$y14^{+}$	1389.66855	1389.6633	-0.0053	253.198013
$b12^{+}$	1396.76174	1396.7627	0.0010	1508.04541
$b26^{2+}$	1398.68844	1398.7573	0.0689	1363.41858
$y15^{+}$	1476.70058	1476.7032	0.0026	357.077484
$b49^{3+}$	1658.12703	1658.1008	-0.0262	656.278015
$y49^{3+}$	1664.13056	1664.1235	-0.0071	457.968933
$b32^{2+}$	1665.78496	1665.8132	0.0282	875.294312
$b15^{+}$	1709.88912	1709.8425	-0.0466	185.520828
$y37^{2+}$	1797.81496	1797.8045	-0.0105	482.84024
$b44^{2+}$	2196.04442	2196.0631	0.0187	412.675201

Table 3. List of TrnB MS/MS fragments.

In the MS/MS spectrum, several fragments containing one to three cysteine residues for each peptide are observed. This will allow a detailed study of TrnC and TrnD in the thioether crosslinks formation once the C-terminal truncation issue is solved. As previously described for other sactipeptides, incubation with iodoacetamide or other cysteine alkylating agents will result in a large mass shift for free thiolates, and no difference in mass if the thioether crosslink is formed[12,23].

In vitro and in vivo attempted studies

Since isolation of TrnA and TrnB by HPLC was not successful, a first attempt to observe thioether crosslink formation was carried out using the final Ni-resin elution (after SUMO protease cleavage), containing TrnA(B) and all the truncated fragments. In anaerobic conditions, after TrnC and TrnD Fe-S clusters reconstitution, TrnC or TrnD, TrnA or TrnB, SAM and sodium dithionite were incubated for 3-4 hours at room temperature. The reaction products were analyzed by mass spectrometry, but the resulting mixture was too complex to be interpreted. None of the expected product or precursor peptide was observed and it was not possible to investigate the nature of the peaks (chromatogram not showed). A large amount of the peptide truncated right before the first cysteine residue was observed in both cases and seemed to be increased, raising again the hypothesis of degradation of the full-length peptide in solution overtime, maybe due to proteases carried over from the initial purification. The experiments have to be repeated after successful isolation of TrnA and TrnB is accomplished.

An attempt to observe the crosslink *in vivo* was also carried out. Co-transformation of pETHSUL-TrnA and pET-28a(+)-TrnC(D) in *E. coli* BL21 Star (DE3) cells was successfully accomplished. Toxicity of SUMO-Trnα (if the crosslinks were formed) was considered, but the cells grew normally and purification of TrnA was carried out. MS analysis resulted in a more complex

mixture compared to the purification of SUMO-TrnA, but the detected peptide did not show any change in mass and the crosslinked one could not be found. Therefore, even if these experiments have to be repeated with more optimal expression conditions, TrnA thioether crosslinks seem not to be formed *in vivo* by either TrnC or TrnD, in aerobic conditions. The only other similar *in vivo* experiment reported was attempted and discussed by Bowers *et al*.[25] for CteA and CteB, without any success.

Conclusions

Thuricin CD is the most potent antibiotic against *C. difficile* known to date. Metronidazole and other drugs currently in use strongly alter the microbiota balance in the gut, with dangerous side effects. Thuricin CD has been shown to have a potency equivalent to those antibiotics, but with far less unspecific activity. It has only been isolated in low amounts from *B. thuringiensis*, and now a larger scale production and advances in clinical trials are needed. For this purpose, understanding of the biosynthetic pathway of this natural product is fundamental. Thuricin CD is a ribosomally synthesized and post-translational modified bacteriocin, formed by two 30-residue subunits, called Trnα and Trnβ. As typical for most RiPPs, thuricin CD is synthesized as two precursor peptides, TrnA and TrnB, which undergo a series of enzymatic reactions, ultimately leading to the mature, bioactive form. Trnα and Trnβ bear the unique cysteine to α-carbon bond, the hallmark of the sactipeptides family. This thioether crosslink is formed through an unknown mechanism by sactisynthases, members of the radical S-adenosylmethionine enzyme superfamily. Sequence alignment show the presence of two genes encoding two putative sactisynthases in thuricin CD operon, called TrnC and TrnD. These enzymes contain the radical SAM motif CXXXCXXC, necessary for coordinating a $[4Fe\text{-}4S]^{2+}$ cluster, and several other cysteine residues in a SPASM domain, near the C-terminus, which may coordinate another Fe-S cluster, as shown for most sactisynthases with similar sequence. After protein expression in iron-enriched media, purification and subsequent *in vitro* reconstitution of the clusters, both enzymes showed the presence of at least one $[4Fe\text{-}4S]^{2+}$ cluster. In order to determine the exact number of clusters, site-directed mutagenesis studies in several different cysteine positions have to be carried out. *In vitro* incubation with iron and sulfur sources could also have resulted in only partial clusters reconstitution. TrnC has shown the ability to reduce and cleave SAM, producing 5'-

deoxyadenosine in the presence of sodium dithionite or titanium(III) citrate. TrnD has not shown the same features. This leads to the conclusion that TrnD Fe-S cluster(s) does not get reduced upon addition of any of the reducing agents tested, or its activity is dependent on other factors. One hypothesis is that addition of the substrate is necessary for activity, or complete reconstitution of multiple clusters is needed. Otherwise, the activity could be contingent on the formation of a TrnD•TrnC complex. EPR spectroscopic analyses will be necessary to provide a better understanding of TrnC and TrnD iron-sulfur clusters.

Expression in *E. coli* and purification of the precursor peptides TrnA and TrnB was complicated, due to their hydrophobicity and small size. Fusion with a Small Ubiquitin-like Modifier protein, with subsequent cleavage using an ubiquitin tertiary structure-specific protease rather than a sequence-specific one, led to successful isolation of the two peptides. Mass spectrometry spectra were acquired, including MS/MS analysis, and the overall result was a very detailed profile of the peptide, which will allow precise detection of the thioether crosslinks. However, TrnA and TrnB are both subjected to multiple truncations from the C-terminus, resulting in a complex mixture of peptides of different size, some of which have been identified. For unknown reasons, the main product of this degradation is a peptide truncated at the residue before the first cysteine, for both TrnA and TrnB. This phenomenon probably occurs during expression, evidenced by the elongated shape of SUMO-TrnA(B) bands on SDS-PAGE. Another more unlikely explanation, could be unspecific activity of the SUMO protease. Spontaneous peptide degradation is also possible, since the full-length peptide peak slowly decreases overtime, and further attempts in HPLC purification failed. Incubation of the degraded peptide mixtures with TrnC and TrnD resulted in MS spectra too complicated to be analyzed. Co-expression of SUMO-TrnA and TrnC(D), with subsequent purification of the peptides, looking for possible *in vivo* crosslinking was also attempted.

Unfortunately, as previously reported for another sactipeptide, this reaction does not seem to occur in aerobic *in vivo* conditions, and no differences in mass in TrnA were detected.

In conclusion, TrnC and TrnD purification procedures have been optimized, and preliminary characterization data have been collected. TrnA and TrnB, precursor peptides to thuricin CD, have been successfully isolated and characterized for the first time. After several attempts in expression and purification of the peptides with different fusion partners, the best results were found using a N-terminal SUMO fusion protein. Optimization of the process is needed, because product degradation is limiting further experiments and analyses. Conditions to initiate TrnC reaction have been established; the crosslinks will be detected by comparison of TrnA and TrnB MS spectra before and after incubation with TrnC. MS/MS fragments containing cysteine residues will allow localization of the crosslinks formed.

Understanding TrnC and TrnD reactions, followed by the assignment of the roles of each gene in thuricin CD gene cluster, will allow the engineering of a production strain, with the final goal of pursuing meaningful studies on its activity against *Clostridium difficile*.

References

1. Lu, W. Antimicrobial peptides. *Semin. Cell Dev. Biol.* **26,** R14–R19 (2018).
2. Babasaki, K., Takao, T., Shimonishi, Y. & Kurahashi, K. Subtilosin Bacillus A , a New Antibiotic subtilis Peptide Produced by Structural and antifungal sub column chromatography on Sephadex. **98,** 585–603 (1985).
3. Arnison, P. G. *et al.* Ribosomally synthesized and post-translationally modified peptide natural products: Overview and recommendations for a universal nomenclature. *Nat. Prod. Rep.* **30,** 108–160 (2013).
4. Repka, L. M., Chekan, J. R., Nair, S. K. & Van Der Donk, W. A. Mechanistic Understanding of Lanthipeptide Biosynthetic Enzymes. *Chem. Rev.* **117,** 5457–5520 (2017).
5. Itoh, H., Matsuoka, S., Kreir, M. & Inoue, M. Design, synthesis and functional analysis of dansylated polytheonamide mimic: An artificial peptide ion channel. *J. Am. Chem. Soc.* **134,** 14011–14018 (2012).
6. Donia, M. S., Ravel, J. & Schmidt, E. W. A global assembly line for cyanobactins. *Nat. Chem. Biol.* **4,** 341–343 (2008).
7. Vassiliadis, G., Destoumieux-Garzón, D., Lombard, C., Rebuffat, S. & Peduzzi, J. Isolation and characterization of two members of the siderophore-microcin family, microcins M and H47. *Antimicrob. Agents Chemother.* **54,** 288–297 (2010).
8. Shimamura, H. *et al.* Structure determination and total synthesis of bottromycin A2: A potent antibiotic against MRSA and VRE. *Angew. Chemie - Int. Ed.* **48,** 914–917 (2009).
9. E Gross and Morell J L. Structure of nisin. *J. Am. Chem. Soc.* **93,** 4634–4635 (1971).
10. Scharf, D. H. *et al.* A dedicated glutathione S -transferase mediates carbon-sulfur bond

formation in gliotoxin biosynthesis. *J. Am. Chem. Soc.* **133,** 12322–12325 (2011).

11. Kawulka, K. *et al.* Structure of subtilosin A, an antimicrobial peptide from Bacillus subtilis with unusual posttranslational modifications linking cysteine sulfurs to α-carbons of phenylalanine and threonine. *J. Am. Chem. Soc.* **125,** 4726–4727 (2003).
12. Flühe, L. *et al.* The radical SAM enzyme AlbA catalyzes thioether bond formation in subtilosin A. *Nat. Chem. Biol.* **8,** 350–357 (2012).
13. Holliday, G. L. *et al. Atlas of the Radical SAM Superfamily: Divergent Evolution of Function Using a "Plug and Play" Domain*. *Methods in Enzymology* **606,** (Elsevier Inc., 2018).
14. Jarrett, J. T. Biotin synthase: Enzyme or reactant? *Chem. Biol.* **12,** 409–410 (2005).
15. Booker, S. J. Unraveling the Pathway of Lipoic Acid Biosynthesis. *Chem. Biol.* **11,** 10–12 (2004).
16. Forouhar, F. *et al.* Two Fe-S clusters catalyze sulfur insertion by radical-SAM methylthiotransferases. *Nat. Chem. Biol.* **9,** 333–338 (2013).
17. Barr, I. *et al.* Demonstration that the radical s-adenosylmethionine (SAM) Enzyme PqqE catalyzes de novo carbon-carbon cross-linking within a peptide substrate PqqA in the presence of the peptide chaperone PqqD. *J. Biol. Chem.* **291,** 8877–8884 (2016).
18. Wieckowski, B. M. *et al.* The PqqD homologous domain of the radical SAM enzyme ThnB is required for thioether bond formation during thurincin H maturation. *FEBS Lett.* **589,** 1802–1806 (2015).
19. Sit, C. S., Van Belkum, M. J., McKay, R. T., Worobo, R. W. & Vederas, J. C. The 3D solution structure of thurincin H, a bacteriocin with four sulfur to α-carbon crosslinks. *Angew. Chemie - Int. Ed.* **50,** 8718–8721 (2011).
20. Rutter, J., Winge, D. R. & Schiffman, J. D. Succinate Dehydrogenase-Assembly, Regulation and Role in Human Disease. *Mitochondrion* **10,** 393–401 (2010).

21. Marahiel, M. A. Two [4Fe-4S] Clusters Containing Radical SAM Enzyme SfkB Catalyze Thioether Bond Formation during the Maturation of the Sporulation Killing Factor. *J. Am. Chem. Soc.* **135,** 959–962 (2013).

22. Grell, T. A. J. *et al.* Structural and spectroscopic analyses of the sporulation killing factor biosynthetic enzyme SkfB, a bacterial AdoMet radical sactisynthase Downloaded from. *J. Biol. Chem.* **14,** (2018).

23. Bruender, N. A., Wilcoxen, J., Britt, R. D. & Bandarian, V. Biochemical and Spectroscopic Characterization of a Radical S-Adenosyl- l -methionine Enzyme Involved in the Formation of a Peptide Thioether Cross-Link. *Biochemistry* **55,** 2122–2134 (2016).

24. Duarte, A. F. de S. *et al.* Hyicin 4244, the first sactibiotic described in staphylococci, exhibits an anti-staphylococcal biofilm activity. *Int. J. Antimicrob. Agents* **51,** 349–356 (2018).

25. Grove, T. L. *et al.* Structural Insights into Thioether Bond Formation in the Biosynthesis of Sactipeptides. *J. Am. Chem. Soc.* **139,** 11734–11744 (2017).

26. Nakai, T. *et al.* The radical S-Adenosyl-L-methionine enzyme QhpD catalyzes sequential formation of intra-protein sulfur-to-methylene carbon thioether bonds. *J. Biol. Chem.* **290,** 11144–11166 (2015).

27. Rea, M. C. *et al.* Thuricin CD, a posttranslationally modified bacteriocin with a narrow spectrum of activity against Clostridium difficile. *Proc. Natl. Acad. Sci.* **107,** 9352–9357 (2010).

28. Chen, L. F. & Sexton, D. J. Clostridium difficile Infection: A Global Perspective of an Epidemic. *Medscape Infect. Dis.* (2008).

29. Guh, A. Y. & Kutty, P. K. *Clostridioides difficile* Infection. *Ann. Intern. Med.* **169,** ITC49 (2018).

30. Centers for Disease Control and Prevention (CDC). *Nearly half a million Americans suffered*

from Clostridium difficile infections in a single year. (2015).

31. Health Protection Surveliance Centre (HPSC). Clostridium difficile-associated disease (CDAD). *Infect. Intest. Dis. Public Heal. Clin. Guid.* (2012).

32. Cotter, P. D., Ross, R. P. & Hill, C. Bacteriocins-a viable alternative to antibiotics? *Nat. Rev. Microbiol.* **11,** 95–105 (2013).

33. Browne, B. L., McClendon, V. & Bedwell, D. M. Mutations within the first LSGGQ motif of Ste6p cause defects in a-factor transport and mating in Saccharomyces cerevisiae. *J. Bacteriol.* **178,** 1712–1719 (1996).

34. Mathur, H., O'Connor, P. M., Cotter, P. D., Hill, C. & Ross, R. P. Heterologous expression of thuricin cd immunity genes in listeria monocytogenes. *Antimicrob. Agents Chemother.* **58,** 3421–3428 (2014).

35. Sit, C. S., McKay, R. T., Hill, C., Ross, R. P., Vederas, J. C. The 3D Structure of Thuricin CD, a Two-Component Bacteriocin with Cysteine Sulfur to α-Carbon Cross-links. *J. Am. Chem. Soc.* **133,** 7680–7683 (2011).

36. Benjdia, A., Guillot, A., Ruffié, P., Leprince, J. & Berteau, O. Post-translational modification of ribosomally synthesized peptides by a radical SAM epimerase in Bacillus subtilis. *Nat. Chem.* **9,** 698–707 (2017).

37. Latham, J. A., Iavarone, A. T., Barr, I., Juthani, P. V. & Klinman, J. P. PqqD is a novel peptide chaperone that forms a ternary complex with the radical S-adenosylmethionine protein PqqE in the pyrroloquinoline quinone biosynthetic pathway. *J. Biol. Chem.* **290,** 12908–12918 (2015).

38. Collins, P. Determination of Iron in Urine Using 4 , 7-Diphenyl- Phosphoryl Chloride Enhancement of Fluorescence and Absorbance of Estrogens in Sulfuric Acid. 1692–1693 doi:10.1021/ac60154a046

39. Beinert, H. Semi-micro methods for analysis of labile sulfide and of labile sulfide plus sulfane sulfur in unusually stable iron-sulfur proteins. *Anal. Biochem.* **131,** 373–378 (1983).

40. Ghabbour, E. A. *Humic Substances - Nature's Most Versatile Materials*. (2004).

41. New England Biolabs Inc. IMPACT™-TWIN - Purification, Ligation and Cyclization of Recombinant Proteins Using Self-Cleavable Affinity Tags. 1–17 (2008).

42. Waugh, D. TEV Protease. 1–7 (2010).

43. Weeks, S. D., Drinker, M. & Loll, P. J. Ligation independent cloning vectors for expression of SUMO fusions. *Protein Expr. Purif.* **53,** 40–50 (2007).

44. Imlay, J. A. Iron-sulphur clusters and the problem with oxygen. *Mol. Microbiol.* **59,** 1073–1082 (2006).

45. Sweeney, W. V & Rabinowitz, J. C. Proteins Containing 4Fe-4S Clusters: An Overview. *Ann. Rev. Biochem.* **49,** 139–61 (1980).

46. Duin, E. C. *et al.* [2Fe-2S] to [4Fe-4S] cluster conversion in Escherichia coli biotin synthase. *Biochemistry* **36,** 11811–11820 (1997).

47. New England BioLabs. Protein Expression & Analysis: pMAL Protein Fusion & Purification System Instruction Manual. 33 (2018).

48. Parent, A. *et al.* The B12-Radical SAM Enzyme PoyC Catalyzes Valine Cβ-Methylation during Polytheonamide Biosynthesis. *J. Am. Chem. Soc.* **138,** 15515–15518 (2016).

49. Li, S. J. & Hochstrasser, M. The Ulp1 SUMO isopeptidase: Distinct domains required for viability, nuclear envelope localization, and substrate specificity. *J. Cell Biol.* **160,** 1069–1081 (2003).

50. Lucigen Corporation. Expresso ™ T7 SUMO Cloning and Expression System. *Lucigen* 1–26 (2016).